Gas-phase structures of large and asymmetric molecules

Graeme R. Kafka

A thesis presented for the degree of
Doctor of Philosophy
University of Edinburgh
2010

Declaration

This thesis has not been submitted, in whole or in part, for any degree at this or any other university. The work is original and my own, carried out under the supervision of Prof. D. W. H. Rankin; where this is not so, credit has been duly given.

Name: Graeme Kafka

Date: 23/11/10

Signed:

Acknowledgements

I would like first to thank my supervisor, David Rankin, for all his help, guidance, and encouragement over the years of my study. Thanks must also go to Sarah Masters, who has been and invaluable help throughout the course of this work and has always been available to discuss aspects of her previous work which laid the foundations for the SEMTEX method.

Also, my thanks go to Heather Robertson for her expertise in collecting data with the electron diffraction machine, and to Derek, Rob, Anthony, Stuart, Phil, Kay and Thomas, all of whom have provided helpful advice at several points along the way.

Thanks must go to the EPSRC National Service for Computational Chemistry Software (NSCCS) for the allocation of computer time, and crucially to the School of Chemistry for funding this work.

A special thanks to those nearest and dearest to me (you know who you are!), who have been a constant and vital support. Finally thanks to anyone else who I've forgotten, who has had to put up with me these past few years – you all deserve a medal!

Abstract

The refinement of large, crowded and asymmetric systems has always proved highly challenging using gas-phase electron diffraction (GED) data. Frequently, local symmetry constraints are still applied to the system to ensure a manageable size of parameter set during the refinement process. This can lead to errors in the heavy-atom structure, as the core atoms move to compensate for the artificially enforced constraints on the peripheral atoms.

This work is focused on the development of the new Structure Enhancement Methodology using Theory and Experiment (SEMTEX) method, which enables the incorporation of high-level *ab initio* data dynamically throughout the refinement process. Initially, the bulky phosphine molecules $OPBu^t_3$ and $HNPBu^t_3$ were studied to test the method, since a ready comparison could be made to previous studies. Subsequent work has seen the technique applied to various different structures. The fluorinated ring compound $C_6F_{11}CF_3$ has been studied, where the peripheral fluorine atoms can be expected to make a large contribution to the overall scattering. The refinement of the structure of the tungsten compound $W(NBu^t)_2(NHBu^t)_2$ followed, presenting a test case for the new method in which more than one conformer was present. The technique was then applied to the family of compounds $C(SiMe_3Cl)_4$, $C(SiMe_3Br)_4$ and $C(SiMe_3F)_4$, large, bulky molecules with more than one type of peripheral atom in the same structure and several conformers present in significant quantity. Finally, the GED refinement of the structure of $Fe_3(CO)_{12}$ is reported, a case where the new method must be further developed in order to be applicable.

Contents

Contents of Supplementary Information on attached CD (Electronic Appendix)

S3. Supplementary information for Chapter Two - The development of the SEMTEX method of total structure determination, and the molecular structures of tri-*tert*-butylphosphine oxide, $OPBu^t_3$, and tri-*tert*-butylphosphine imide, $HNPBu^t_3$, using the new method.

S6a. Supplementary Information for Chapter Six - Comparison of the gas-phase molecular structure of $C(SiMe_2Cl)_4$ determined using each of the SARACEN and SEMTEX methods.

S6b. Supplementary Information for Chapter Six - Comparison of the gas-phase molecular structure of $C(SiMe_2F)_4$ determined using each of the SARACEN and SEMTEX methods.

Table S6b.1. Comparison of calculated parameters for $C(SiMe_2F)_4$ calculated using the MM3 and MP2 methods.

Table S6b.2. Interatomic distances (r_a/pm) and amplitudes of vibration (u/pm) from the SARACEN GED refinement of the structure of $C(SiMe_2F)_4$.

Table S6b.3. Final gas electron diffraction coordinates from the SARACEN refinement of $C(SiMe_2F)_4$.

Table S6b.4. Correlation matrix for the SARACEN refinement of $C(SiMe_2F)_4$.

Table S6b.5. Interatomic distances (r_a/pm) and amplitudes of vibration (u/pm) from the SEMTEX GED refinement of the structure of $C(SiMe_2F)_4$.

Table S6b.6. Final gas electron diffraction coordinates from the SEMTEX refinement of $C(SiMe_2F)_4$.

Table S6b.7. Correlation matrix for the SEMTEX refinement of $C(SiMe_2F)_4$.

Table S6b.8. Refined peripheral-atom parameters for $C(SiMe_2F)_4$ calculated using the SEMTEX method.

S6c. Supplementary Information for Chapter Six - Comparison of the gas-phase molecular structure of $C(SiMe_2Br)_4$ determined using each of the SARACEN and SEMTEX methods.

Table S6c.1. Comparison of calculated parameters for $C(SiMe_2Br)_4$ calculated using the MM3 and MP2 methods.

Table S6c.2. Interatomic distances (r_a/pm) and amplitudes of vibration (u/pm) from the SARACEN GED refinement of the structure of $C(SiMe_2Br)_4$.

Table S6c.3. Final gas electron diffraction coordinates from the SARACEN refinement of $C(SiMe_2Br)_4$.

Table S6c.4. Correlation matrix for the SARACEN refinement of $C(SiMe_2Br)_4$.

Table S6c.5. Interatomic distances (r_a/pm) and amplitudes of vibration (u/pm) from the SEMTEX GED refinement of the structure of $C(SiMe_2Br)_4$.

Table S6c.6. Final gas electron diffraction coordinates from the SEMTEX refinement of $C(SiMe_2Br)_4$.

Table S6c.7. Correlation matrix for the SEMTEX refinement of $C(SiMe_2Br)_4$.

Table S6c.8. Refined peripheral-atom parameters for $C(SiMe_2Br)_4$ calculated using the SEMTEX method.

S7. The gas-phase structure of $Fe_3[CO]_{12}$

Table S7.1. Interatomic distances (r_a/pm) and amplitudes of vibration (u/pm) from the SARACEN GED refinement of the structure of $Fe_3[CO]_{12}$.

Table S7.2. Final gas electron diffraction coordinates from the SARACEN refinement of $Fe_3[CO]_{12}$.

Table S7.3. Correlation matrix for the SARACEN refinement of $Fe_3[CO]_{12}$.

List of terms, symbols and abbreviations used

$\angle$	angle (in degrees)
Å	Angstrom
av.	average parameter
B	Becke's exchange functional (1988)
B3	Becke's three-parameter hybrid functional
Bu^t	tertiary butyl group
DFT	Density Functional Theory
diff.	difference parameter
DYNAMITE	Dynamic Interaction of Theory and Experiment (method)
E	energy
e.s.d.	estimated standard deviation
EPSRC	Engineering and Physical Sciences Research Council
eV	electron volt
ϕ	torsional angle
f_x	scattering factor for atom X
GED	gas-phase electron diffraction
GTO	Gaussian-type orbital
$\hat{H}$	Hamiltonian operator
HF	Hartree-Fock theory
IR	infrared spectroscopy
k	perpendicular correction factor
K	degrees Kelvin
kV	kilovolt
LANL2DZ	basis set using Los Alamos effective core potential on heavy atoms
LYP	correlation function of Lee, Yang and Parr
m	metre
Me	methyl group
mm	millimetre
MM	molecular mechanics
MOCED	Molecular Orbital Constrained Electron Diffraction (method)

MP2	second-order Møller-Plesset perturbation series
nm	nanometre
NMR	Nuclear Magnetic Resonance (spectroscopy)
NSCCS	National Service for Chemistry Computing Software
PES	potential energy surface
pm	picometre
PW91	exchange and correlation functional of Pardew and Wang
r	interatomic distance
R_G	goodness-of-fit parameter (including correlation between parameters)
RHF	spin-restricted Hartree-Fock
RMS	root mean square
s	second
SARACEN	Structural analysis restrained by *ab initio* calculations for electron diffraction (method)
SHRINK	Program used to determine curvilinear vibrational corrections
T	temperature
tert	tertiary
u	amplitude of vibration
Ψ	wavefunction

1. Introduction to gas-phase electron diffraction

1.1 General introduction

The physical and chemical properties exhibited by compounds are governed in large part by the size and shape of the molecules. Accordingly, the study of molecular structure is of paramount importance in the advancement of chemical knowledge. The ability to study molecular structures in the gas phase accurately is particularly important, as here they are free from intermolecular interactions, which can affect the structures of molecules in the solid state, allowing us to obtain unbiased and "intrinsic" structural information. Futhermore, *ab initio* calculations generally provide us with structural geometries of free molecules, and therefore comparison between theory and experiment can easily be made where a reliable gas-phase experimental structure can be obtained.

Molecular structure has traditionally been studied using the standard techniques of diffraction and spectroscopy, tailored to the phase in which the molecule is to be studied. For example, X-ray diffraction is the pre-eminent technique for the study of crystal structures, while in the gas phase, electron diffraction and rotational spectroscopy are the commonly-used methods. The usefulness of rotational spectroscopy is, however, generally limited to small molecules, in order that sufficient structural information can be gleaned from the data without the need for an unfeasible number of isotopic substitutions. Gas-phase electron diffraction (GED) therefore plays a unique and key role in the determination of molecular structures of larger molecules, such as the ones studied in this work.

The recent emergence of powerful theoretical methods for the prediction of molecular structure has provided another valuable tool for the structural chemist. Fuelled by advances in available computational power to perform the calculations, both *ab initio* and Density Functional Theory (DFT) methods are now suitable for application to the range of molecules generally studied using GED. When used in combination with experimental data, these results of such calculations can aid greatly in the complete refinement of the structures of more challenging molecules. Conversely, accurate experimental structures are indispensable in order to calibrate and further refine the new theoretical methods constantly under development.

1.2 Early development of the electron diffraction technique

The roots of gas-phase electron diffraction as an experimental technique can be traced back to the early development of the theory of quantum mechanics. The first fundamentally key development was the classic double-slit experiment of Young in 1801,[1] which demonstrated the formation of an interference pattern when an incident light wave is diffracted from multiple sources. In Young's experiment, a cylindrical wavefront produced by diffraction of a light wave through a narrow slit is incident upon a screen containing two parallel slits. The diffraction occurring from these produces two coherent wavefronts, which if detected at a subsequent point will produce a pattern of light and dark areas, corresponding to constructive and destructive interference.

Secondly, in 1924 de Broglie hypothesised[2] that all moving particles have a wavelength associated with them, the length of which is inversely proportional to their momentum. Therefore, an electron can be considered both as a particle and as a wave; and, owing to its extremely small mass, it will have a wavelength of measurable magnitude. This indicated that a beam of electrons would be capable of producing a diffraction pattern such as that observed by Young, if incident on a material from which scattering occurred.

The experimental confirmation of the de Broglie hypothesis was provided in 1927 by two independent Nobel prize-winning experiments, one by Davisson and Germer,[3] and the other by George Thomson.[4] Davisson and Germer fired a beam of electrons at a nickel crystal, and observed the that the scattered electron intensity showed an angular dependence that would give the same diffraction pattern as predicted for X-rays. Thomson, meanwhile, demonstrated that diffraction occurred when a beam of electrons passed through a thin gold foil.

1.3 The electron diffraction experiment

The principle behind the electron diffraction experiment is based on these observations, and its development as an experimental technique dates back to 1930, when the first electron diffraction experiment was reported by Mark and Wierl.[5]

A beam of incident electrons, of energy between 20 and 100 keV, is fired from an electron gun at a sample in a vacuum chamber ('the diffraction chamber'). When such a beam of electrons is incident on a molecule, diffraction occurs from each pair of atoms, which act like a pair of slits in Young's experiment. The diffracted electrons then interfere, since the wavelength of the electrons is comparable to the separation between the atoms in a molecule, and the resulting interference pattern is recorded on a photographic film, which is used as the detector. The sample molecules are injected into the diffraction chamber by means of a nozzle system, which can be designed to suit many specific experimental requirements, as outlined later. The chamber must be kept at high vacuum in order to prevent the electrons from being scattered by sources other than the target molecules. A schematic of the generic GED experiment is shown in Figure 1.1, and an example of the recorded interference pattern is given in Figure 1.2.

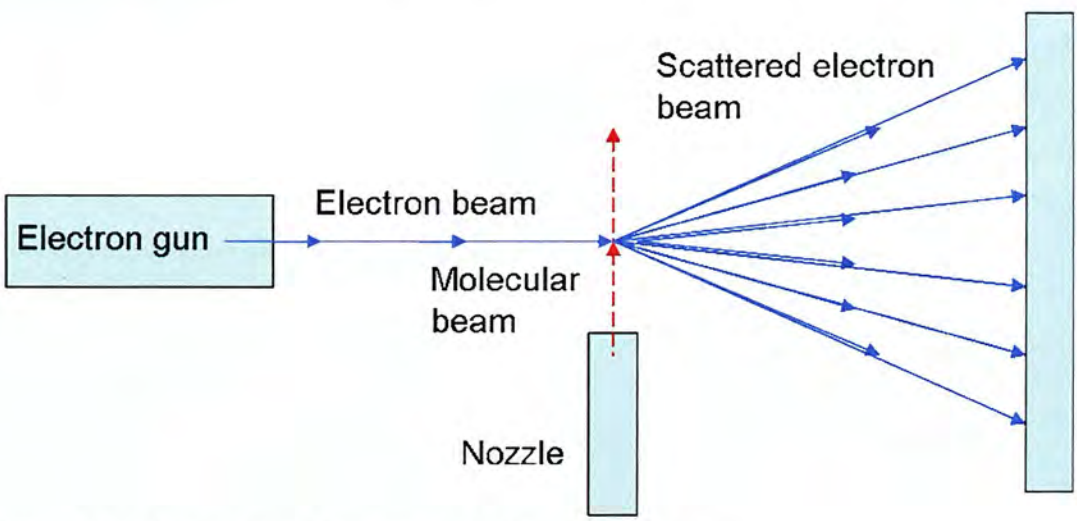


Figure 1.1. Schematic of the generic GED experiment.

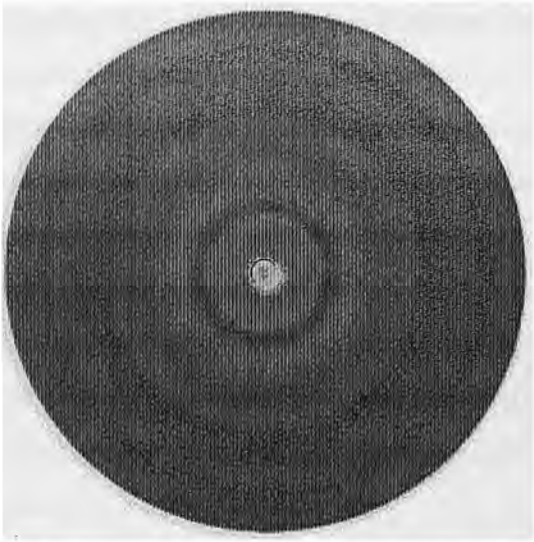


Figure 1.2. Developed photographic plate containing data from an electron diffraction experiment.

As can be seen in Figure 1.2, the diffraction pattern consists of a series of concentric rings at different distances r from the centre of the plate, caused by the summation of the interference from all the different, randomly-orientated molecules in the gas sample. The optical densities are then measured as a function of the scattering angle. However, as the scattering angle is dependent on the wavelength of the electrons in the beam, the data are instead interpreted in terms of s units, independent of the wavelength of the electrons, and defined through the simple relation

$$s = \frac{4\pi \sin\left(\frac{\theta}{2}\right)}{\lambda} \qquad \textbf{Equation 1}$$

This has the desirable effect of making the data independent of the apparatus on which they are collected, ensuring compatibility across the different refinement programs used by various research groups. Generally, two different sample-to-plate distances will be used during data collection for each sample. A longer distance provides better information on scattering at low s values, and a shorter one is able to measure scattering out to much larger s values. In this way, better-quality data can be obtained across the whole range of s.

A rapid drop-off in scattering intensities occurs with increasing scattering angle (and thus distance from the centre of the plate): the intensity falls off as $1/s^4$.[6] This causes a very uneven intensity distribution to reach the plate, meaning that large-angle scattering is impossible to discern against the very intense low-angle pattern. To combat this, a specially-designed piece of metal, known as a rotating sector, is introduced just before the photographic plate.[7]

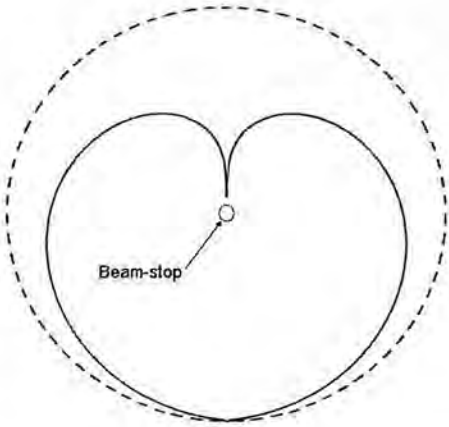


Figure 1.3. Rotating sector used in the Edinburgh GED apparatus.

The rotating sector is a rapidly-rotating metallic disc with an opening cut to the fourth power of the scattering angle, which thus compensates for the intensity drop-off by increasing the relative exposure time of the plate to high scattering angles, preventing the information on large-angle scattering from being lost. Even with this in place, the workable range for which intensity measurements can be carried out is generally only up to $s \sim 30$ nm^{-1}. It is also necessary to prevent possible back-scattering by the use of a beam-stop (a small cylinder) in the centre of the rotating sector to collect unscattered electrons, unfortunately making it impossible to collect data at very small scattering angles.

The total experimental intensity is a function of the scattering angle, s, and can be viewed as the sum of contributions from both molecular scattering, atomic scattering and from the 'experimental background':

$$I_t(s) = I_m(s) + I_a(s) + I_b(s) \qquad \textbf{Equation 2}$$

The experimental background contains extraneous (background) scattering from various sources. Once the total scattering has been obtained, the atomic scattering is calculated by simply summing the contributions from each atom in the molecule, and then removed in order to isolate the important information – that relating to molecular properties. This gives the total experimental molecular-scattering intensity curve (MIC) plus the extraneous background, which is empirically determined and eliminated during the subsequent structure analysis. The experimental MIC contains scattering from every pair of atoms in the molecule, and also from every combination of three atoms, four atoms and so on. In most cases, such multiple-atom scattering makes a minimal contribution to the overall intensity, and can therefore be discounted in the interpretation of the data. However, in some cases, the three-atom scattering contribution cannot be readily discarded and must be included explicitly in the scattering factor calculations.[8]

When approximated as the sum of the contributions from all of the pairs of atoms in the molecule, the total molecular scattering intensity takes the form

$$I_{molecular} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{f_i f_j}{s^4} \cos\left[\eta_i(s) - \eta_j(s)\right] \exp\left(-\frac{1}{2}u_{ij}^2 s^2\right) \frac{\sin\left[s\left(r_{ij} - \kappa_{ij}s^2\right)\right]}{s r_{ij}}$$

Equation 3

where $i \neq j$.

The MIC thus takes the form of a damped sine curve, as can be seen in Figure 1.4. The implication of some of the other terms in this equation will be discussed subsequently.

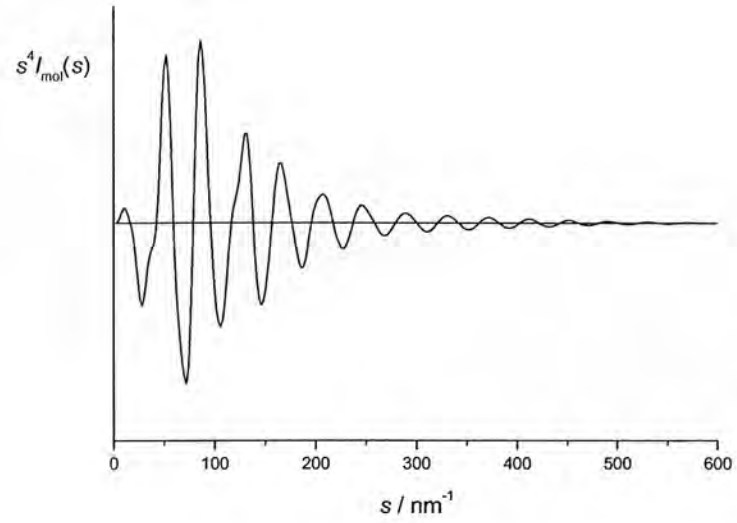


Figure 1.4. Example of a molecular-scattering intensity curve. The intensity can be seen to drop off with increasing scattering angle s.

The experimental molecular-scattering intensity curve is compared with a theoretical one, calculated for the predicted molecular model geometry. This model is constructed by specifying the positions of every atom in the molecule in terms of bond lengths, bond angles and torsional angles. If the model structure is not known, theoretical distributions for various models may have to be compared with the experimental one in order to establish which best fits the data.

A Fourier transformation of the molecular-scattering intensity curve is then performed to give the radial distribution curve (RDC):

$$f(r) = \int_0^{s_{max}} sM(s)\exp(-as^2)\sin(sr)\mathrm{d}s \qquad \textbf{Equation 4}$$

This is much more easily interpreted than the molecular intensity curve. A sample curve is shown in Figure 1.5. The absence of any experimental data at very small scattering angles can cause a shift in the zero line of the RDC, and theoretical data are therefore introduced into the scattering intensity curves between $s = 0$ and $s = s_{min}$ to avoid this.

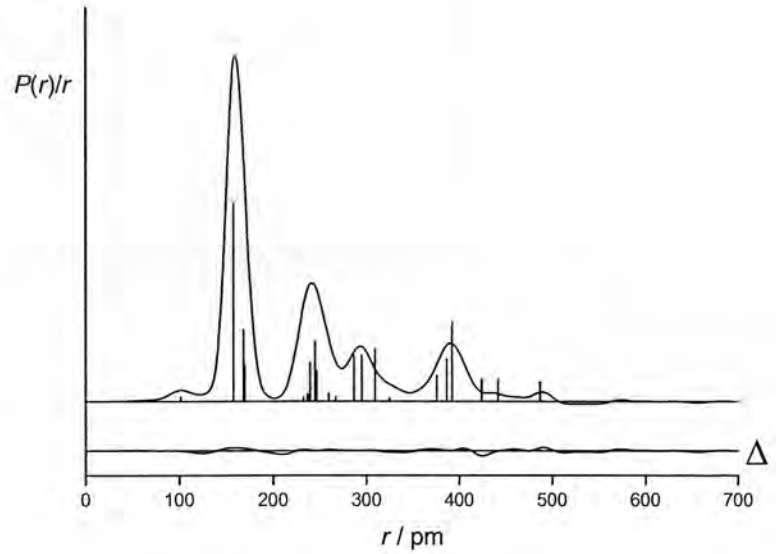


Figure 1.5. Example of a radial distribution curve. The differences between experimental and theoretical models are illustrated in the difference curve Δ (beneath the RDC).

As can be seen, the curve consists of a series of peaks at different values of r (interatomic distances.) Each of these peaks corresponds to the distance between a pair of atoms in the molecule. Thus, the electron diffraction experiment gives information on the distances between pairs of atoms in the molecule, both bonded and non-bonded.

The structural parameters are then refined using a least-squares method, based on the molecular-scattering intensities, using a well-established program.[9] Bond lengths, angles and torsion angles are used to describe the structure, in accordance with the construction of the model geometry. For an N-atom molecule, $(3N - 6)$ parameters are required to describe its geometry fully. Fortunately, the number of required parameters can be significantly reduced by taking advantage of any global or local symmetry possessed by the molecule. The application of artificial symmetry constraints, however, can have serious consequences for the accuracy of the molecular structure obtained, as will be seen later. Caution must therefore be exercised when writing the molecular model that any symmetry elements included are genuine properties of the structure. Care must also be taken to ensure that the starting (theoretical model) parameters used for the least-squares refinement are not

far from their real values: the model geometry must be close to the true case. Thus, a good initial agreement between the experimental and theoretical values is required.

The "goodness-of-fit" parameter, or R factor, is used to quantify the agreement between the molecular geometry of the refining structure, and the experimental molecular-scattering intensity data. An R factor of < 10% (R_G < 0.010) is generally taken to signify a reliable molecular structure determination.

1.4 Limitations of standard GED

1.4.1 Difficulty in obtaining information on light atoms

When studying molecules with many similar interatomic distances, such distances will frequently appear underneath the same peak in the radial distribution curve. When one intermolecular distance contributes much more to the overall scattering intensity than the others, the more intense one will refine well; the less intense ones will not. This can make it difficult to determine the complete molecular structure accurately, especially in cases when several distances lie under the same peak. This is the case for all of the refinements carried out in this thesis.

The relative areas of the peaks corresponding to each separate interatomic distance can be calculated using Equation 5:

$$Area \propto \frac{Z_i Z_j n_{ij}}{r_{ij}} \qquad \textbf{Equation 5}$$

where r_{ij} is the distance between nuclei i and j, Z_i and Z_j are their atomic numbers, and n_{ij} the number of times this distance occurs. Thus it can be seen that, for example, where the atomic numbers are similar for two different pairs, only the distance that occurs more often in the molecule will refine well. Furthermore, distances involving heavy atoms will refine much better than those involving light ones that lie under the same peak. Inclusion of additional data from other sources is

therefore required in order to refine the complete structure accurately, and this process will be discussed later in this chapter.

1.4.2. Molecular vibrations

Another challenge to the structural accuracy attainable using GED arises as a result of molecular vibrations. One effect of molecular vibration on the measured interatomic distance for a simple linear triatomic molecule like CO_2 is shown in Figure 1.6.

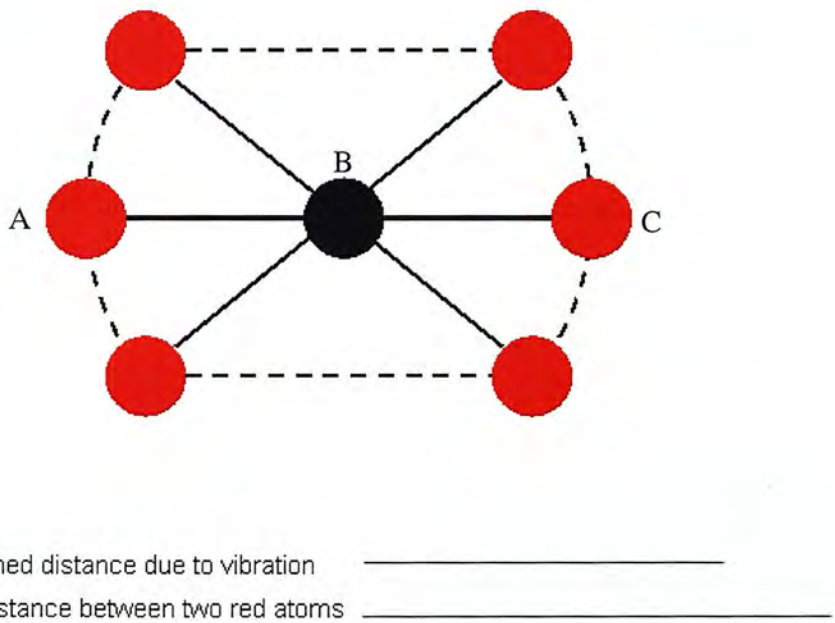


Figure 1.6. Vibration of a linear triatomic molecule (Adapted from Ref. 10)

As can be seen, the molecule spends almost its entire time bent away from the linear geometry, and thus the distance between the two terminal atoms A and C is clearly, on average, less than twice the distance from A to B. Since the measured scattering intensity is a result of the contribution of millions of scattered electrons, with each 'seeing' the molecule more or less instantaneously, this will be borne out in the distance data obtained from the experiment. It would therefore seem from measurement of these two distances that the molecule was in fact bent, rather than linear, which is obviously incorrect. Such a phenomenon is also found for non-linear

systems, known as the shrinkage effect,[11-12] and corrections for vibrational effects must be made in order to account for this.

In this work, the SHRINK[13-14] program has been used to determine corrections to the vibrational model. This works by taking the force constants obtained from calculated force fields to calculate curvilinear corrections, modelling the curved motion of the atoms as they bend away from the linear geometry as seen in Figure 1.6. This enables a good estimate of how molecular vibrations have influenced interatomic distances in the electron diffraction data.

The distances obtained directly from the electron diffraction experiment are the reciprocal of the averaged reciprocal distances, denoted r_a:

$$r_a = \langle r^{-1} \rangle^{-1} \qquad \textbf{Equation 6}$$

To obtain the average internuclear separation distance, denoted r_g, the amplitude of vibration between the two atoms must be taken into consideration:

$$r_g = \langle r \rangle = r_a + \frac{u^2}{r_e} \qquad \textbf{Equation 7}$$

Ideally, we would like to obtain the equilibrium geometry, with interatomic distances denoted r_e, which is defined as the value of r at the overall minimum on the potential energy in a hypothetical vibrationless state. This is desirable for two reasons. First, since such distances are independent of the method used to obtain them, structural information from different experimental techniques can be directly compared if the value of r_e can be obtained in each case. Secondly, the results of *ab initio* calculations yield equilibrium structures, and as will be discussed later, such calculations are extremely useful in providing additional information to help overcome many of the traditional limitations of standard GED. It is therefore helpful to have distances from the refinement that are as close to the r_e values as possible.

In this work, the SHRINK program has been used to provide distance corrections using a harmonic force field derived from the minimum on the potential-energy surface. These corrections are denoted k_{h1}, and yield r_{h1} distances, where

$$r_{h1} = r_a + \frac{u^2}{r} - k_{h1}$$ **Equation 8**

As long as the molecular structure under investigation is not too far from the bottom of the potential well (and so the distances are not affected significantly by the anharmonicity of the potential energy curve), these corrections provide a good estimate of the effects of shrinkage whilst not being too time-consuming to calculate. It can also be seen from Equation 3 that the molecular intensity includes an exponential term, which drops off as u^2, where u is the amplitude of vibration between a pair of atoms. This means that any interatomic distance with a large vibrational amplitude will contribute very little to the MIC and hence will be poorly defined using GED. This makes the study of molecules containing hydrogen bonds and other weak interactions difficult.

1.4.3 Phase effects

As an electron propagates through the field of an atom, it speeds up by an amount dependent on the atom's charge. This causes a corresponding shortening of the de Broglie wavelength of the electron. Upon leaving the atomic field, the wavelength lengthens again as the electron slows back to its original speed, but its phase has been been altered by the temporary shortening of the wavelength. A pictorial representation of this phenomenon is given in Figure 1.7.

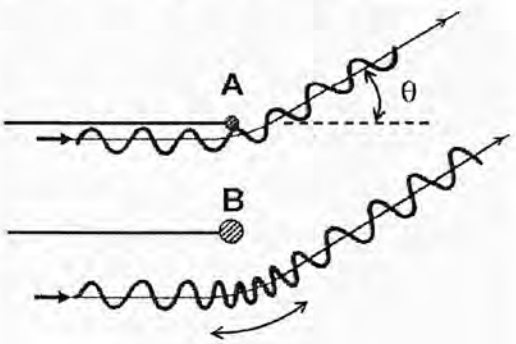


Figure 1.7. Representation of the phase shift experienced by electrons as they propagate through the atomic fields of A and B, a light and a heavy atom, respectively. (Adapted from Ref. 15)

It can be seen that Equation 3 contains a cosine phase term, $\cos[\eta_i(s) - \eta_j(s)]$. Clearly, in the case where atoms i and j have similar atomic numbers, they will induce a very similar change in phase and so $\eta_i(s) - \eta_j(s) \approx 0$, and this phase term will have a negligible effect on the overall scattering intensity for all values of s. However, when the two atoms have very different atomic numbers, this will not be true and in certain regions of s the measured scattering intensities will be significantly reduced. This has the effect of causing the interatomic distance A-B in the radial distribution curve to appear as a split peak instead of a single well-defined maximum. The theoretical curves make allowance for this effect by using complex scattering factors to calculate the predicted molecular scattering. However, in some cases this phase effect can cause the calculated R factor for the refinement to be much higher, which can mislead the experimentalist as to the quality of the refinement.

1.5 Theoretical calculations

The electron diffraction experiment gives information on the internuclear distances in gas-phase molecules. Thus its results are particularly well suited to comparison with those from theoretical calculations, which give structures in terms of sets of nuclear coordinates. These calculations are on individual molecules, and can therefore also be considered to be on gas-phase structures. Such calculations can thus be combined with the experimental data to provide important additional information, which can help overcome many of the inherent limitations of the basic GED experiment.

With the rapid increase in available computational power in recent years, great advancements have been made in the development of computational methods for structure determination. Such methods are now routinely used to complement experimental results, considerably increasing the range of molecular structures that can be determined using methods such as GED. A brief description of those theoretical methods used in this work follows.

1.5.1 Molecular mechanics

One of the quickest and simplest methods of obtaining a theoretical molecular structure is through the use of molecular mechanics (MM).[16] This method can be simply viewed as treating each atom as a "ball" or single particle, with bonded (and some non-bonded) distances treated as harmonic oscillators or "springs", and uses Newtonian mechanics to calculate the properties of the structure from this basic model. The total potential energy of the molecule is calculated through the use of a force field, a well-defined set of parameters containing information relating to the bond lengths, angles and dihedrals in the molecule, along with electrostatic and van der Waals energy terms. Various different force fields are in common use, the MM3 force field being used in this work. [17-19]

Harmonic potentials are used to calculate the energy components from bond lengths and bond angles. For example, in the case of bonded distances, each is assigned an equilibrium bond length, and a force constant k, mimicking the degree of vibration of this bonded distance. The bond energy is thus defined in Newtonian terms by the degree of compressibility of the "spring" for each bond. Since the van der Waals term is modelled using a Lennard-Jones potential, it drops off rapidly with increasing distance, and thus only the van der Waals interactions with close neighbours need be considered for each atom, saving a lot of computational time. A cut-off radius is therefore introduced, with the interactions between any pair of atoms further apart than this radius deemed to be zero. In the case of the electrostatic potential, however, the drop-off only scales as r^{-1} (as it takes the form of the Coulomb potential), and is therefore much more difficult to model accurately using MM.

The parameters for each MM force field are generally derived from the results of trusted experiments, and also other theoretical methods, such as *ab initio*, which will be discussed presently. MM provides a computationally cheap way to model molecular structures theoretically, and this will be seen to be useful to GED structure refinement in subsequent discussion. However, its reliance on pre-determined parameters and approximation of the model using Newtonian mechanics make it perform more poorly in general than more in-depth treatments, which treat the molecule using quantum mechanics, and these will now be discussed.

1.5.2 *Ab initio* calculations

Ab initio ("from the beginning") molecular orbital theory is a heavily used and powerful method of theoretical structure determination. This calculates molecular properties from first principles alone, by calculating a solution to the time-independent Schrödinger Equation (SE)

$$E\Psi = \hat{H}\Psi \qquad \textbf{Equation 9}$$

for the molecule under investigation. The SE cannot be solved exactly for systems with more than one electron, however, and an approximate solution must be sought by making simplifications to the wavefunction Ψ and Hamiltonian operator $\hat{H}$. The Hamiltonian comprises the potential energies of nuclear-nuclear repulsion, electron-electron repulsion and nuclear-electronic attraction, along with the kinetic energies of the nuclei and electrons.

The most well-known simplification to the Hamiltonian is the Born-Oppenheimer approximation.[19-20] This assumes that as the electrons move so much faster than the nuclei in a molecule, the nuclei can be considered as stationary, allowing separation of the nuclear and electronic wavefunctions. Hence, the nuclear-nuclear repulsion term can be considered to be a constant, and the nuclear kinetic energy term can be approximated as zero.

The most basic starting point for *ab initio* calculations is the use of the Hartree-Fock potential to approximate the electron-electron repulsion term. This assumes that the term can be approximated by treating each individual electron as moving in a uniform electronic field generated by all the other electrons in the molecule. This allows the intractable multi-electron Schrödinger Equation to be replaced by a set of single-electron SEs, for which exact solutions can be found. Simple use of the Hartree-Fock method provides a good starting-point for theoretical calculations, reliably accounting for around 99% of the energy in a molecule. It does not account for all of the energy in a molecule, however, since it neglects the effect of electron correlation, which is present in any system containing more than one electron. This leads to calculated bond distances being too short, especially in molecules with areas

of high electron density such as double bonds or lone pairs of electrons. The desire to include the effects of electron correlation has led to the development of several post-HF methods.

Most post-Hartree-Fock *ab initio* methods operate by adding extra terms to the HF wavefunction to approximate the effects of electron correlation. One such method is the commonly-used Møller-Plesset perturbation series,[21] which accounts for such effects by introducing perturbations to the HF wavefunction. In this work, the method used is the MP2 level of theory (representing the fact that perturbations up to second-order are applied to the wavefunction). This method represents a considerable increase in accuracy over the basic HF method for most systems, whilst not being too computationally demanding.

1.5.3 Density Functional Theory

An alternative theoretical method for determining molecular structures is provided by Density Functional Theory (DFT). This is based on the fact that electron density of a system can be used to make a complete determination of its ground-state energy.[22] Since the electron density at any point in the system is defined by only three co-ordinates, this method is much less computationally expensive than the *ab initio* methods previously discussed.

DFT methods still require approximations to be made, however, in this case to the functional relating electron density to the energy of the electrons. In pure DFT, therefore, both electron exchange and correlation effects are approximated, in contrast to HF, which calculates exactly the exchange term but neglects correlation completely.

Development of functionals for use in DFT is carried out by using reliable experimental data to fit parameters for the functional, and then testing against large sets of atoms and molecules.

In order to include the effects of exchange more accurately, hybrid DFT methods can be used. These combine the exchange energy calculated exactly using HF with the approximate correlation energy from DFT. The method commonly used in this thesis

is B3PW91.[23,24] This comprises a method of including exact exchange energy, the B3 functional, and a correlation functional, PW91.

1.5.4 Basis sets

To find an approximate solution to the Schrödinger Equation, the wavefunction Ψ must also be simplified. Since the nuclei are already assumed to be fixed in space by the use of the Born-Oppenheimer approximation, the wavefunction need only provide a description of the atomic orbitals, describing the region of space through which the electrons move. The general way of doing this is by using Gaussian functions to model the orbitals (GTOs).

In order to reproduce the atomic orbitals to an acceptable accuracy, it is necessary to use a combination of multiple GTOs. The set of GTOs applied to each atom in a molecule is known as a basis set. In theory, the larger the basis set (the more Gaussian functions used), the better the replication of the atomic orbitals would be, and thus the more accurate the wavefunction. However, computational constraints again mean that the series of GTOs must be limited.

An example of a simple basis set that has been used frequently in this work for initial calculations on structures is 6-31G.[25] The numbers refer to the number of GTOs describing each atomic orbital. This is a split-valence basis set, with the 6 referring to the core electrons, and the 3 and 1 referring to the inner and outer valence electrons. A split-valence basis set has the advantage of allowing the use more basis functions to describe the more chemically important valence electrons, while treating the core electrons with a minimal basis set to ensure computational feasibility. Such a basis set is described as being "single-zeta" with respect to the core electrons, where six GTOs comprising a single basis function are used to describe the atomic orbitals of each electron. It is "double-zeta" with respect to the valence electrons, where two basis functions are used for each electron, one comprised of three GTO and the other a single GTO.

To compensate for the fact that some atoms can change significantly in size and shape upon becoming part of a molecule, two additional functions can be used. Polarisation functions, denoted by a *, add functions of higher angular momentum

than would be normal in the ground state. Diffuse functions, denoted by a +, allow the orbitals to spread out and occupy more space, and are important in systems with areas of high electron density such as double-bonds and lone pairs.

Such basis sets are not applicable to all atoms. In general, larger atoms often need to be treated by approximating the core electrons (which have a negligible effect on bonding) by a pseudopotential, and handling the valence electrons in the same way as for smaller atoms. This has the effect of greatly speeding up calculations for such atoms.

1.6 Combining theory and experiment

1.6.1 The SARACEN method

As discussed previously, one of the major limitations of electron diffraction traditionally lay in the difficulty of resolving different interatomic distances whose peaks in the radial distribution curve overlap. These then appear as one broad peak, under which several interatomic distances may lie. Without additional data, there is often insufficient information to determine all of the individual distances which lie under such peaks.

Accordingly, attempts were made to combat this problem. First, the MOCED (Molecular Orbital Constrained Electron Diffraction) method[26] was developed, allowing difficult-to-refine parameters to be constrained to calculated values to enable refinement of the complete structure. Naturally, it is desirable that such troublesome parameters should be allowed some flexibility, as otherwise artificial constraints are imposed upon any other refining parameters that are correlated with these. The development of the SARACEN (Structure Analysis Restrained by *Ab initio* Calculations for Electron diffractioN) method[27-29] at the University of Edinburgh was a large step forward in overcoming this problem. This method involves the use of computed *ab initio* data as flexible restraints in the refinement process, allowing all parameters to be refined. This has allowed study of a far wider range of structures using electron diffraction than was previously possible.

Whilst SARACEN was a revolutionary development in the field of GED, it is still intrinsically a heavy-atom method – symmetry constraints often still have to be applied to the light atoms to ensure a manageable number of refining parameters. This of course is not ideal. For example, although the contribution to the total scattering from an individual hydrogen atom is very small, the presence of many such atoms in a structure can make a large overall scattering contribution, as illustrated in Figure 1.8.

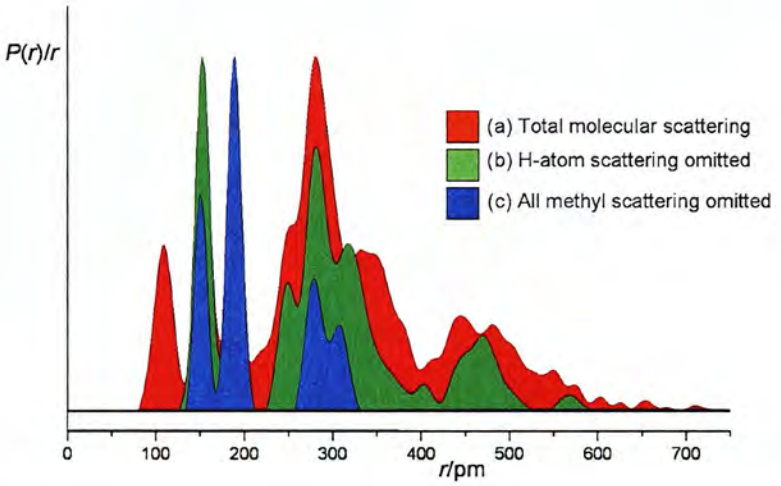


Figure 1.8. An illustration of the contributions of various molecular components to the total scattering intensity (here for $OPBu^t_3$). Comparison of the red (total) and green (without hydrogen) scattering intensities clearly shows the importance of modelling light-atom scattering correctly. (Adapted from Ref. 30)

In sterically crowded molecules, interactions with other atoms may cause these light atoms to be displaced, and the effect of these displacements on the overall scattering cannot be ignored. If the hydrogen atoms in a particular group are constrained to a certain symmetry, this can cause error in the heavy-atom structure: the heavier atoms are not constrained and thus move to compensate as a consequence of the restrictions on the light atoms, as illustrated by Figure 1.9.

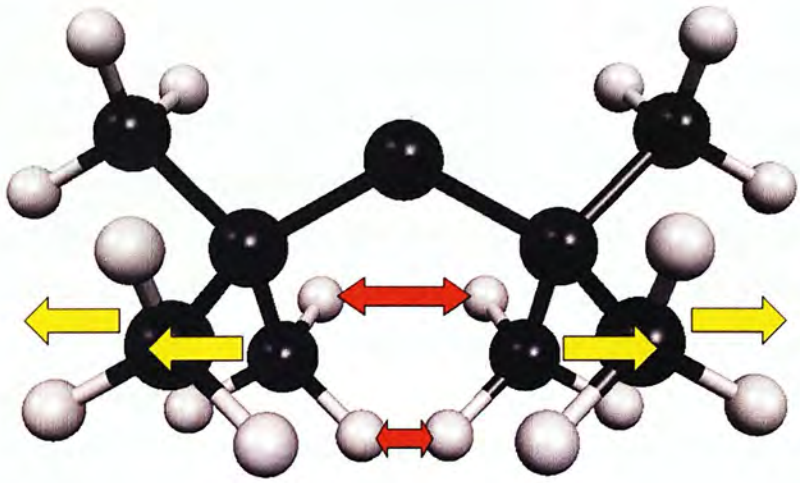


Figure 1.9. Example of the effect of symmetry constraints on the refined molecular structure of a sterically crowded molecule. The H atoms (red arrows) would like to move apart, but are constrained by symmetry. This forces the methyl carbon atoms (yellow arrows) to move instead, affecting the heavy-atom structure.(Adapted from Ref. 30)

Therefore, a new method was required in order to eliminate symmetry constraints on the light atoms and provide a better model of asymmetric structural features in crowded molecules.

1.6.2 The DYNAMITE method

In order to overcome the problem of assumptions of local symmetry in light-atom groups, the DYNAMITE (DYNAMic Interaction of Theory and Experiment) method was developed.[30] In DYNAMITE, an inexpensive computational method is used to continually update the light-atom positions in each refinement cycle throughout the process. This removes local symmetry assumptions made by SARACEN and allows light-atom groups to adopt asymmetric conformations if required. The computational method used in the work done so far was molecular mechanics, using the TINKER package with the MM3 force field.

The MM method is called once for each refining parameter, for each cycle of refinement. The heavy-atom structure is still determined from the refinement of the

GED data. With the heavy-atom positions fixed at the values determined by GED, the MM code is used to calculate the positions of all the light atoms. This then returns a set of light-atom coordinates, which are read back into the least-squares refinement program, removing the need for symmetry of light-atom groups and thus preventing the heavy-atom positions from being artificially influenced by this requirement.

The MM3 method does not calculate the absolute values for the bond lengths, angles and torsions relating to the light atoms very well. As a result, the values it returns must be scaled back to the average GED parameter from the continuing refinement in order to fit the experimental data.

The use of DYNAMITE on several large, sterically crowded molecules has been found to give a better fit to experimental data than the corresponding SARACEN refinements, in which local symmetry constraints still had to be applied to the light atoms in order to maintain a manageable number of refining parameters. This clearly indicates that removal of the local symmetry constraints for the light atoms is important. Significant differences in parameters relating to heavy atoms were observed upon implementation of the DYNAMITE code for compounds including $OPBu^t_3$ and $HNPBu^t_3$.[30,31] Equally important was the test for an uncrowded system $[Sn(C_2P_2Bu^t_2)]$.[32] In this case, implementation of DYNAMITE resulted in no change to the heavy-atom structure, as would be expected. Thus DYNAMITE is not artificially imposing a structure and any change in the heavy-atom skeleton is a real effect of the removal of symmetry constraints.

1.6.3 Beyond DYNAMITE - the SEMTEX method

Whilst DYNAMITE was a large step forward in the accurate structure determination of large, sterically crowded and asymmetric molecules, it still contains room for improvement. The use of molecular mechanics, a low-level theoretical method, as the computational model for the light-atom positions is clearly not optimal. In the interests of accuracy, it would be desirable to use a high-level method to calculate the light-atom positions during each refinement cycle. Such a method can be *ab initio* or DFT, depending upon the molecule under investigation. These have the advantage of

being significantly more accurate than molecular mechanics simulations, which are still based on a set of previously-determined parameters. Unfortunately, it is not feasible at the moment to include such calculations directly in place of the MM calculations in DYNAMITE, due to the computationally demanding nature of these methods. Therefore, the higher-level theoretical method must be indirectly implemented in the refinement process. In order to achieve this, the method of SEMTEX (Structural Enhancement Methodology using Theory and EXperiment) has been devised, as follows.

First, the heavy atoms in the structure are fixed at their GED positions, and both a molecular mechanics and an *ab initio*/DFT calculation are performed on this structure to find the positions of the light atoms in each case. (In this work, MP2 calculations are used where an *ab initio* method is deemed preferable). From these, sets of bond lengths, angles and torsions can be calculated for each case to describe the light-atom positions in relation to the heavy-atom skeleton.

A set of differences is then calculated between the light-atom parameters of the MP2/DFT calculation and those of the molecular mechanics one. Subsequently, in each refinement cycle, the light-atom parameters returned by the MM3 code (the "DYNAMITE part") are then modified by this set of differences, to account for the difference between the parameters returned by the two separate levels of theory.

During the progression of the refinement, the heavy-atom parameters will change in value somewhat. This causes the initial theoretical calculations, and thus the differences between the two sets of light-atom parameters, to become inaccurate. Therefore, it is necessary to update the theoretical difference set periodically by performing further *ab initio* calculations on the updated heavy-atom positions, in order to maintain accuracy. If this is not done, the light-atom parameters begin to "drift", that is, to continually change in one direction with each cycle of refinement until they become unfeasible. The inclusion of repeated *ab initio* calculations naturally increases the refinement time greatly, but not to the unworkable degree which an *in situ ab initio* method would do. The work of this thesis is concerned primarily with the development and testing of this new method, and its application to the refinement of various structures that have proven challenging for the standard SARACEN method.

1.7 References

1. T. Young, *Philos. Trans. R. Soc. London, Ser. A,* 1802, **92**, 12.
2. L. de Broglie, *Philos. Mag.*, 1924, **47**, 446.
3. C. J. Davisson and L. H. Germer, *Nature* (*London*), 1927, **119**, 558.
4. G .P. Thomson, *Nature* (*London*), 1927, **120**, 802.
5. H. Mark and R. Wierl, *Z. Elektrochem.*, 1930, **36**, 675.
6. J. Tremmel and I. Hargittai, Gas Electron Diffraction Experiment, in *Stereochemical Applications of Gas-Phase Electron Diffraction, Part A: The Electron Diffraction Technique*, ed. I. Hargittai and M. Hargittai, VCH, 1988.
7. P. Debye, *Phys. Z.*, 1939, **40**, 404.
8. P. D. McCaffrey, J. K. Dewhurst, D. W. H. Rankin, R. J. Mawhorter and S. Sharma. *J. Chem. Phys.*, 2008, **128**, 204304.
9. S. L. Hinchley, H. E. Robertson, K. B. Borisenko, A. R. Turner, B. F. Johnston, D. W. H. Rankin, M. Ahmadian, J. N. Jones and A. H. Cowley, *Dalton Trans.*, 2004, 2469.
10. T. Foerster, *Ph.D. Thesis*, University of Edinburgh, 2006.
11. O. Bastiansen and M. Trætteberg, *Acta. Crystallogr.*, 1960, **13**, 1108.
12. R. Stølevik, H. M. Seip and S. J. Cyvin, *Chem. Phys. Lett.*, 1972, **15**, 263.
13. V. A. Sipachev, *J. Mol. Struct.* (*THEOCHEM*), 1985, **121**, 143.
14. V. A. Sipachev, *J. Mol. Struct.*, 2001, **567**, 67.
15. B. F. Johnston, *Ph.D. Thesis*, University of Edinburgh, 2002.
16. C. A. Morrison, *Advanced Experimental and Computational Methods,* Chemistry 4/5 Lecture Course, University of Edinburgh, 2004.
16. N. L. Allinger, Y. H. Yuh and J-H. Lii, Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 1. *J. Am. Chem. Soc.*, 1989, **111**, 8551.
17. J-H. Lii and N. L. Allinger, Molecular Mechanics, The MM3 Force Field for Hydrocarbons. 2. Vibrational Frequencies and Thermodynamics, *J. Am. Chem. Soc.*, 1989, **111**, 8566.
18. J-H. Lii and N. L. Allinger, Molecular Mechanics, The MM3 Force Field for Hydrocarbons. 3. The van der Waals Potentials and Crystal data for Aliphatic and Aromatic Hydrocarbons, *J. Am. Chem. Soc.*, **111**, 8576.
19. M. Born and J. R. Oppenheimer, *Ann. Phys.* (*Berlin*) 1927, **84**, 457.

20. B. T. Sutcliffe, *Adv. Quantum Chem.*, 1997, **28**, 65.
21. C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.
22. W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.
23. J. P. Perdew and Y. Wang, *Phys. Rev. B*: *Condens. Matter*, 1992, **45**, 13244.
24. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
25. M. Francl, W. Pietro, W. Hehre, J. S. Binkley, M. Gordon, D. DeFrees and J. A. Pople, *J. Chem. Phys*, 1982, **77(7)**, 3654.
26. N. S. Chiu, H. L. Sellers, L. Schaefer and K. Kohata, *J. Am. Chem. Soc.* **1979**, *101*, 5883.
27. P. T. Brain, C. A. Morrison, S. Parsons and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.* **1996**, 4589.
28. A. J. Blake, P. T. Brain, H. McNab, J. Miller, C. A. Morrison, S. Parsons, D. W. H. Rankin, H. E. Robertson and B. A. Smart, *J. Phys. Chem.* **1996**, *100*, 12280.
29. N. W. Mitzel and D. W. H. Rankin, *Dalton Trans.* **2003**, 3650.
30. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Dalton Trans*. 2004, 384.
31. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Inorg. Chem.*, 2004, **43**, 5522.
32. D. A. Wann, S. L. Hinchley, K. B. Borisenko, H. E. Robertson, M. D. Francis, J. F. Nixon and D. W. H. Rankin, *Dalton Trans.*, 2005, 1972.

2. Experimental details

2.1. General overview

As outlined in the previous chapter, the molecular structures reported in this thesis were determined using a combination of computational methods and experimental data from GED. This chapter contains details of the experimental method, which are broadly identical across all of the separate refinements, except where otherwise stated in the relevant chapter. The GED data was collected using the University of Edinburgh electron diffraction apparatus, as described below.

The first stage is to obtain an optimised molecular geometry from theoretical calculations. This is then used to provide starting values of parameters in the model used in the refinement of the GED data. Once all the desired parameters and amplitudes of vibration have been allowed to refine according to the SARACEN method, the SEMTEX code is activated and the structure refined again, with the peripheral-atom positions now determined computationally.

2.2. Computational methods

Throughout the work in this thesis, all potential-energy surface searches, frequency calculations and geometry optimisations were performed on the Columbus cluster, maintained by the National Service for Computational Chemistry Software (NSCCS),[1] using the Gaussian 03 program.[2] Our implementation of the DYNAMITE optimisation method uses the TINKER molecular mechanics package with the MM3 parameter set.[3]

The first step in each analysis was to conduct a search of the potential-energy surface using the RHF method and the 3-21G*[4-6] basis set on all atoms, in order to identify all possible conformers of the molecule. Each conformer located was optimised at the HF/6-31G*[7,8] level, and when multiple conformers were found their relative energies were calculated. The amount of each conformer in the experimental mixture was then predicted using the Boltzmann distribution, and conformers predicted to comprise < 5% of the experimental mixture were not included in the electron diffraction model. Higher-level geometry optimisations were then carried out on the included conformers. Where an *ab initio* method was deemed preferable, the MP2[9]

level of theory was used, while the B3PW91[10,11] method was applied where DFT calculations were used. Split-valence triple-zeta Pople-style basis sets were used in each case. The lowest-energy structures were then used to provide the initial parameter values for the electron diffraction model. The method and basis set used for each molecule can be found in Table 2.1.

Table 2.1. Levels of theory used for geometry optimisations of each molecule studied in this thesis. Both the method and the basis set used are specificed in each case.

Molecule	*ab initio* method	DFT method
$OPBu^t_3$	MP2/6-311G*[12,13]	-
$HNPBu^t_3$	MP2/6-311G*	-
$C_6F_{11}CF_3$	MP2/6-311+G*[14]	-
$W(NBu^t)_2(NHBu^t)_2$	MP2/6-311G*	B3PW91/6-311+G*
$C(SiMe_2Cl)_4$	MP2/6-311G*	-
$C(SiMe_2F)_4$	MP2/6-311G*	-
$C(SiMe_2Br)_4$	MP2/6-311G*	-
$Fe_3(CO)_{12}$	MP2/6-311G*	B3PW91/6-311G*

The force fields were obtained using analytic second derivatives of the energy with respect to nuclear coordinates calculated at the HF/6-31G* level for each molecule. These were then used by the program SHRINK[15,16] to provide estimates of the amplitudes of vibration (u_{h1}) and curvilinear vibrational correction factors (k_{h1}) to distances required for the GED refinements. The force fields were also used to calculate the frequencies for the optimised structures. In every case, all calculated frequencies were real, indicating that each structure represented a minimum on the global potential-energy surface for that molecule.

2.3. Electron diffraction data

The Edinburgh gas-phase electron diffraction (GED) apparatus was used to collect data. An accelerating voltage of 40 keV was used (wavelength *ca.* 6.0 pm). Scattering intensities were recorded on Kodak Electron Image film at two nozzle-to-film distances for each sample, to maximise the scattering angle over which data were collected. Multiple films were recorded at each nozzle-to-film distance. In order to obtain suitable vapour pressures the samples and inlet nozzle were heated. In the refinements of $OPBut_3$ and $HNPBut_3$, old electron diffraction data were reintroduced directly into the refinement program, as these structures were used as test cases for the new method, for direct comparison with previous work.[17,18] The specific temperatures and nozzle-to-film distances used for data collection of each of the other molecules are given in Table 2.2.

Table 2.2. Nozzle-to-film distances and experimental temperatures used in the GED data collection for each molecule. "Short" and "Long" refer to the short and long nozzle-to-film distances, respectively. All distances are in mm and temperatures in K.

Molecule	Short dist.	Long dist.	Sample temperature		Nozzle temperature	
			Short	Long	Short	Long
$C_6F_{11}CF_3$	128.1	292.1	298	293	298	298
$W(NBu^t)_2(NHBu^t)_2$	97.4	260.8	443	423	450	429
$C(SiMe_2Cl)_4$	100.1	255.6	464	435	477	454
$C(SiMe_2F)_4$	97.8	260.3	435	399	446	414
$C(SiMe_2Br)_4$	94.8	259.2	496	477	506	496
$Fe_3(CO)_{12}$	91.6	254.6	410	372	425	392

The photographic films were scanned using an Epson Expression 1680 Pro flatbed scanner. The data reduction and least-squares refinement were carried out using the ed@ed program,[19] with scattering factors developed in-house.[20]

2.4. SEMTEX methodology

Initially, the heavy-atom structure of the molecule being studied is determined *via* a standard SARACEN[21-23] GED refinement, i.e. restrained where necessary by parameters calculated *ab initio*. The Cartesian coordinates of this refined structure are output to a data file. Geometry optimisations at both the MP2/DFT and molecular mechanics levels of theory are then performed, with the heavy-atom coordinates fixed to their refined GED values in each case. This returns two sets of peripheral-atom coordinates, whose differences are attributable entirely to the effects of the differing levels of theory used. The position of each SEMTEX-active peripheral atom in each computed structure is then derived in terms of a bond length, angle and dihedral angle. By subtracting the MP2 values of these three parameters from the molecular mechanics ones, a set of difference parameters is obtained. These are then introduced as additional data in the refinement process, modifying the parameters returned continually by the MM code to reflect the differences between the two levels of theory. The absolute differences between peripheral-atom parameters are thus derived *ab initio*, but they are modified dynamically at the MM level.

As the entire structure refines using this method, the initial heavy-atom structure (as calculated by the SARACEN method) will change over several cycles of refinement. This renders the initial set of differences obsolete, as they were calculated based on heavy-atom coordinates frozen at values that are no longer valid. As a consequence, it is necessary to repeat the theoretical calculations periodically and recalculate the difference set during the refinement. The method works with any two types of calculation, an expensive, high-level one, performed few times, and a low-level, cheap one, performed repeatedly.

2.5. References

1. EPSRC National Service for Computational Chemistry Software. URL: http://www.nsccs.ac.uk.

2. Gaussian 03, Revision C. 01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian Inc., Wallingford CT, 2004.

3. J. W. Ponder, *TINKER© molecular mechanics freeware package*, Version 3.8, October 2000, St. Louis, MO.

4. J. S. Binkle, J. A. Pople and W. J. Hehre, *J. Am. Chem. Soc.* 1980, 102, 939.

5. M. S. Gordon, J. S. Binkley, J. A. Pople, J. W. Pietro, and W. J. Hehre, *J. Am. Chem. Soc.* 1982, **104**, 2797.

6. W. J. Pietro, M. M. Francl, W. J. Hehre, D. J. DeFrees, J. A. Pople and J. S. Binkley, *J. Am. Chem. Soc.* 1982, **104**, 5039.

7. W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.* 1972, **56**, 2257.

8. P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* 1973, **28**, 213.

9. C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.

10. J. P. Perdew and Y. Wang, *Phys. Rev. B*: *Condens. Matter*, 1992, **45**, 13244.

11. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.

12. A. D. McLean and G. S. Chandler, *J. Chem. Phys.* 1980, **72**, 5639.

13. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.* 1980, **72**, 650.

14. M. S. Gordon, *Chem. Phys. Lett.* 1980, **76**, 163.

15. V. A. Sipachev, *J. Mol. Struct.* (*THEOCHEM*), 1985, **121**, 143.

16. V. A. Sipachev, *J. Mol. Struct.*, 2001, **567**, 67.

17. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Dalton Trans*. 2004, 384.

18. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Inorg. Chem.*, 2004, **43**, 5522.

19. S. L. Hinchley, H. E. Robertson, K. B. Borisenko, A. R. Turner, B. F. Johnston, D. W. H. Rankin, M. Ahmadian, J. N. Jones and A. H. Cowley, *Dalton Trans.* 2004, 2469.

20. P. D. McCaffrey, *Ph.D. Thesis*, University of Edinburgh, 2007.

21. P. T. Brain, C. A. Morrison, S. Parsons and D. W. H. Rankin, *J. Chem. Soc., Dalton Trans.* **1996**, 4589.

22. A. J. Blake, P. T. Brain, H. McNab, J. Miller, C. A. Morrison, S. Parsons, D. W. H. Rankin, H. E. Robertson and B. A. Smart, *J. Phys. Chem.* **1996**, *100*, 12280.

23. N. W. Mitzel and D. W. H. Rankin, *Dalton Trans.* **2003**, 3650.

3. The development of the SEMTEX method of total structure determination, and the molecular structures of tri-*tert*-butylphosphine oxide, $OPBu^t_3$, and tri-*tert*-butylphosphine imide, $HNPBu^t_3$, using the new method

3.1. Introduction

The work in this chapter is concerned with the initial development and testing of the SEMTEX (Structure Enhancement Methodology for Theory and EXperiment) method, as described in the preceding chapters. The results of the first two structure determinations carried out using this new method are reported here.

The structures of both tri-*tert*-butylphosphine oxide ($OPBu^t_3$) and tri-*tert*-butylphosphine imide ($HNPBu^t_3$) have previously been investigated using the SARACEN and DYNAMITE methods.[1,2] In this work, these refinements have been re-evaluated and the results used as starting geometries for structural analyses using the new SEMTEX method. They have been chosen so that maximum insight can be gained into the workings of the new method.

For this work, slight improvements to the scaling routines in the DYNAMITE code have been implemented. Small scaling errors, previously undetectably small, were revealed by the wider range of values produced during the SEMTEX studies.

3.2. Experimental

3.2.1. Computational methods

For both the molecules under investigation, only one conformer was located. Geometry optimisations and amplitudes of vibration calculations were carried out as detailed in Chapter 2. The lowest energy structures are shown in Figures 3.1 and 3.2.

3.2.2. Electron diffraction data

Original digital molecular-intensity scattering intensity data for both $OPBu^t_3$ and $HNPBu^t_3$[3] were reintroduced directly into the ed@ed Edinburgh electron diffraction refinement program without further modification. The scattering factors of Ross *et al.*[4] were used in the refinements.

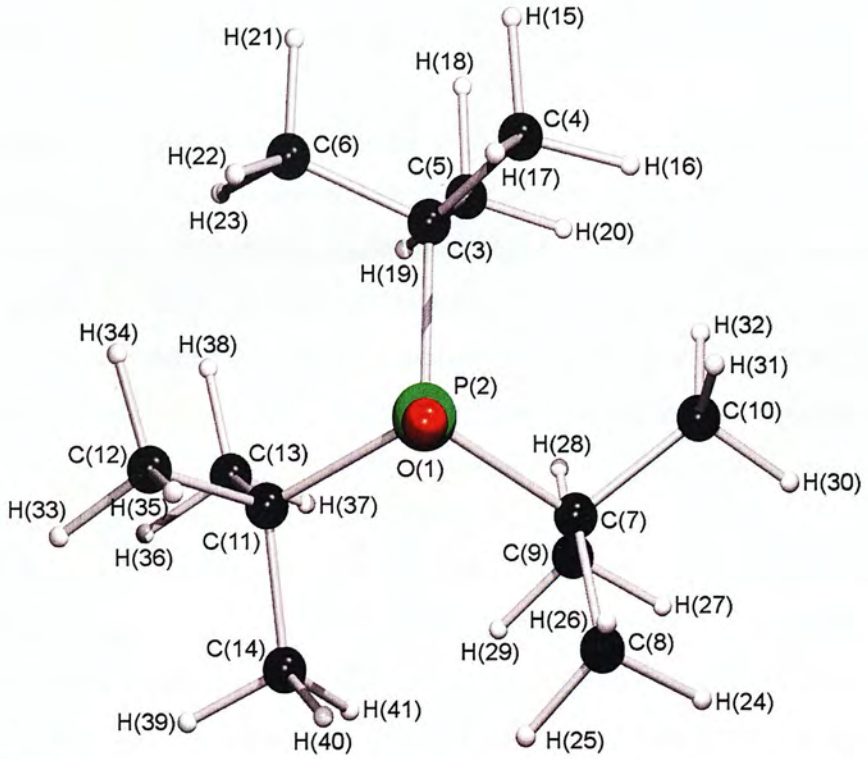


Figure 3.1. Gas-phase molecular structure of $OPBu^t_3$ viewed along the O-P bond (C_3 rotation axis).

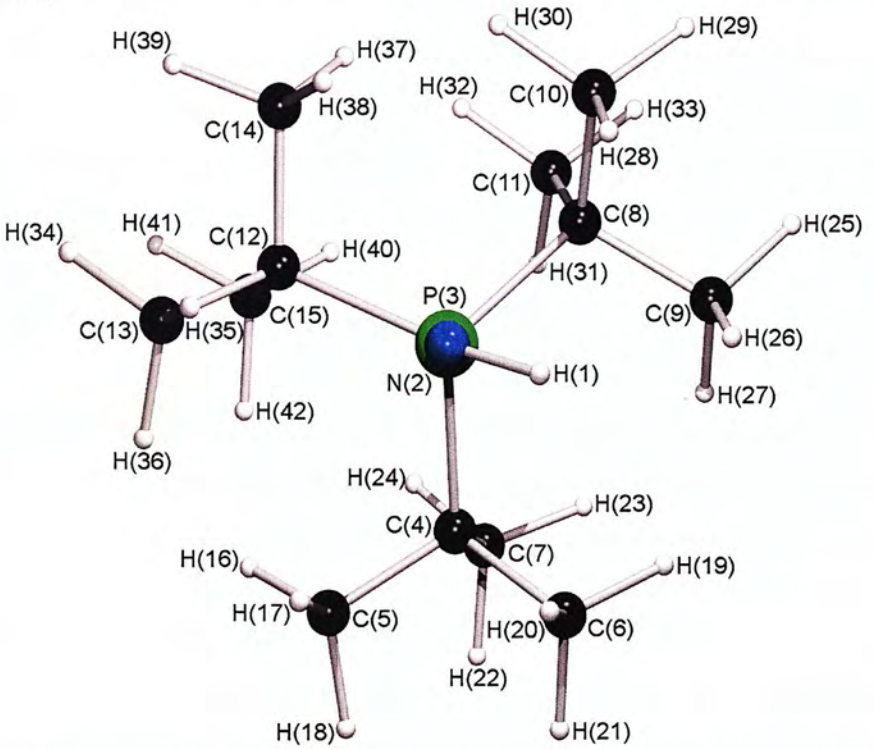


Figure 3.2. Gas-phase molecular structure of $HNPBu^t_3$ viewed along the N-P bond. In this case, the molecule has C_1 symmetry.

3.2.3. Electron diffraction model

3.2.3.1. Tri-tert-butylphosphine oxide

The structure was defined using a model with C_3 symmetry, as indicated by the *ab initio* calculations described above. Fifteen independent geometric parameters were required, comprising four bond lengths, seven bond angles and differences and four torsion parameters. These can be found in Table 3.1.

The heavy-atom bond lengths were described by *r*C-C, *r*P-C and *r*P-O ($p_{1\text{-}3}$). *Ab initio* calculations showed that while in principle there are three different C-C distances in a *tert*-butyl group, the differences are insignificantly small, and so there is no benefit in using more than one parameter.

Independent heavy-atom bond-angle parameters were ∠O-P-C (p_4), an average and two difference parameters [P(2)-C(3)-C(4) – P(2)-C(3)-C(5) and P(2)-C(3)-C(4) – P(2)-C(3)-C(6)] to describe the P-C-C angles ($p_{5\text{-}7}$), and two C-C-C angles, C(4)-C(3)-C(5) and C(4)-C(3)-C(6) ($p_{8,9}$). An angle describing the torsion of the *tert*-butyl group around the P-C bond was also included (p_{10}).

Five parameters were also included to describe the starting positions of the hydrogen atoms. These comprised *r*C-H (p_{11}), ∠C-C-H (p_{12}), and three parameters to describe the torsions of the three methyl groups about their respective C-C bonds ($p_{13\text{-}15}$). In the DYNAMITE and SEMTEX refinements, p_{11} and p_{12} represented the mean values of those for all nine hydrogen atoms in a *tert*-butyl group.

3.2.3.2. Tri-tert-butylphosphine imide

The structure was defined using a model of C_1 symmetry. Altogether forty-two independent geometric parameters were required to describe the structure, comprising seven bond lengths and differences, twenty-two bond angles and differences and thirteen torsion parameters. These can be found in Table 3.2.

The bond lengths were described by the N-H distance (p_1), the average and difference of the C-C and P-N bond lengths ($p_{2\text{-}3}$), and the average and two differences for the P-C bond lengths ($p_{4\text{-}6}$). Independent bond-angle parameters included three average and difference parameters for the P-C-C angles of each *tert*-

butyl group ($p_{7\text{-}9,12\text{-}14,17\text{-}19}$), with two associated C-C-C angles ($p_{10,11,15,16,20,21}$). The P-N-H angle was also used (p_{27}), as were two C-P-C angles ($p_{25,26}$), and the average and two difference parameters to describe the N-P-C angles ($p_{22\text{-}24}$). Torsional parameters were three tert-butyl group torsions, ϕN(2)-P(3)-C(4)-C(5), ϕN(2)-P(3)-C(8)-C(9) and ϕN(2)-P(3)-C(12)-C(13) ($p_{28\text{-}30}$), and ϕH-N-P-C(4) (p_{31}).

Finally, eleven parameters were included to describe the starting positions of the peripheral hydrogen atoms. These comprised the mean C-H bond length (p_{32}), mean ∠C-C-H bond angle (p_{33}), and nine parameters to describe the torsions of the three methyl groups in each butyl group about their adjacent C-C bonds ($p_{34\text{-}42}$). Again, in the DYNAMITE and SEMTEX refinements, p_{32} and p_{33} represented the mean values of those for all nine hydrogen atoms in a *tert*-butyl group.

Table 3.1. Refined and calculated parameters for $OPBu^t_3$ (distances in pm, angles in °) for the SARACEN, DYNAMITE and SEMTEX studies.

no.	parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_1	rC-C	153.8	154.0(2)	153.9(2)	154.0(2)	
p_2	rC-P	189.0	189.1(3)	189.0(2)	188.7(2)	
p_3	rP-O	151.0	149.9(5)	149.8(5)	149.8(5)	149.8(6)
p_4	∠O-P-C	109.5	107.6(3)	107.6(2)	107.7(2)	
p_5	∠P-C-C av.	111.0	111.6(2)	111.6(2)	111.9(1)	
p_6	∠P(2)-C(3)-C(4) − ∠P(2)-C(3)-C(5)	-8.4	-9.6(8)	-10.3(7)	-10.7(7)	
p_7	∠P(2)-C(3)-C(4) − ∠P(2)-C(3)-C(6)	-1.7	-1.8(7)	-1.9(7)	-2.1(7)	-1.7(10)
p_8	∠C(4)-C(3)-C(5)	109.6	111.9(7)	111.8(5)	111.4(7)	
p_9	∠C(4)-C(3)-C(6)	105.1	107.2(4)	107.2(4)	107.2(4)	

no.	parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{10}	ϕO-P-C-C(4)	40.5	36.7(8)	36.2(7)	36.2(7)	
p_{11}	rC-H	109.2	108.1(2)	108.0(2)	108.0(2)	
p_{12}	$\angle$C-C-H	111.0	111.0(6)	111.7(4)	112.2(4)	
p_{13}	ϕP-C-C(4)-H(15)	-169.3	-167.1(24)	-167.4(24)	-167.7(24)	-168.9(26)
p_{14}	ϕP-C-C(5)-H(18)	176.7	176.6(25)	176.7(27)	176.2(26)	176.6(26)
p_{15}	ϕP-C-C(6)-H(21)	169.1	168.9(24)	169.4(24)	169.1(24)	169.3(27)

Table 3.2. Refined and calculated parameters for $HNPBu^t_3$ (distances in pm, angles in °) for the SARACEN, DYNAMITE and SEMTEX studies.

no.	Parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_1	rN-H	101.6	101.1(5)	101.1(6)	101.1(5)	101.0(5)

no.	Parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_2	*r*C-C/P-N av.	154.2	155.9(2)	156.0(2)	155.9(2)	
p_3	*r*C-C/P-N diff	5.0	5.4(5)	5.4(5)	5.4(5)	5.0(5)
p_4	*r*P-C av.	189.8	190.4(4)	191.2(4)	191.1(5)	
p_5	*r*P-C d1	-2.4	-2.4(5)	-2.4(6)	-2.5(5)	-2.4(5)
p_6	*r*P-C d2	-2.5	-2.7(5)	-2.5(6)	-2.5(5)	-2.7(5)
p_7	∠P-C-C av. (gp 1)	110.9	110.9(8)	111.2(9)	111.1(9)	110.9(10)
p_8	∠P-C-C d1 (gp 1)	0.6	-0.2(10)	-0.4(10)	-0.4(10)	0.6(10)
p_9	∠P-C-C d2 (gp 1)	-6.1	-5.5(10)	-5.1(10)	-5.0(10)	-6.0(10)
p_{10}	∠C(5)-C(4)-C(6)	108.4	108.1(10)	107.9(11)	107.9(11)	108.4(10)
p_{11}	∠C(5)-C(4)-C(7)	105.5	105.1(10)	105.6(10)	105.7(10)	105.4(10)
p_{12}	∠P-C-C av. (gp 2)	111.1	112.5(10)	111.7(10)	111.6(10)	

no.	Parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{13}	∠P-C-C d1 (gp 2)	0.9	1.6(10)	1.1(11)	1.2(11)	0.9(10)
p_{14}	∠P-C-C d2 (gp 2)	-8.3	-8.3(10)	-8.0(10)	-8.1(10)	-8.3(10)
p_{15}	∠C(9)-C(8)-C(10)	104.7	104.6(10	104.8(11)	104.9(11)	104.7(10)
p_{16}	∠C(9)-C(8)-C(11)	108.5	108.5(10)	108.5(11)	108.6(11)	108.6(10)
p_{17}	∠P-C-C av. (gp 3)	111.2	109.6(11)	106.7(9)	106.5(9)	
p_{18}	∠P-C-C d1 (gp 3)	-1.4	-1.0(10)	-1.5(11)	-1.5(11)	-1.3(10)
p_{19}	∠P-C-C d2 (gp 3)	-8.2	-8.4(10)	-8.7(10)	-8.8(10)	-8.1(10)
p_{20}	∠C(13)-C(12)-C(14)	105.2	105.4(10)	106.3(10)	106.4(10)	105.2(10)
p_{21}	∠C(13)-C(12)-C(15)	109.1	109.6(10)	110.2(11)	110.3(11)	109.1(10)
p_{22}	∠N-P-C av.	106.5	107.3(3)	107.2(3)	107.2(3)	

no.	Parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{23}	∠N-P-C d1	10.3	9.9(13)	11.8(14)	11.8(15)	10.3(15)
p_{24}	∠N-P-C d2	0.8	0.7(15)	-0.6(15)	-0.5(15)	0.8(15)
p_{25}	∠C(4)-P(3)-C(8)	109.7	109.7(9)	110.5(9)	110.6(9)	109.7(10)
p_{26}	∠C(4)-P(3)-C(12)	109.8	110.1(8)	109.8(8)	109.7(8)	109.8(10)
p_{27}	∠P-N-H	115.7	115.8(11)	115.9(11)	115.9(11)	115.7(10)
p_{28}	ϕN(2)-P(3)-C(4)-C(5)	73.8	74.3(21)	72.1(21)	72.3(20)	73.8(25)
p_{29}	ϕN(2)-P(3)-C(8)-C(9)	67.6	67.2(19)	66.4(20)	66.8(21)	67.6(25)
p_{30}	ϕN(2)-P(3)-C(12)-C(13)	-36.3	-34.7(20)	-39.3(21)	-39.6(20)	-36.3(25)
p_{31}	ϕH-N-P-C	-173.0	-173.0(11)	-173.0(11)	-173.0(11)	-172.9(10)
p_{32}	*r*C-H	109.2	114.5(3)	114.9(3)	115.0(3)	

no.	Parameter	MP2/6-311G* (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{33}	∠C-C-H	107.6	108.6(8)	110.4(7)	110.5(8)	
p_{34}	ϕP(3)-C(4)-C(5)-H(16)	72.6	70.9(26)	71.4(27)	71.4(27)	72.6(25)
p_{35}	ϕP(3)-C(4)-C(6)-H(19)	-70.3	176.4(27)	176.3(27)	176.5(27)	175.9(25)
p_{36}	ϕP(3)-C(4)-C(7)-H(22)	-175.9	69.4(26)	69.4(27)	69.7(27)	70.2(25)
p_{37}	ϕP(3)-C(8)-C(9)-H(25)	73.8	73.1(26)	73.5(28)	73.7(28)	73.8(25)
p_{38}	ϕP(3)-C(8)-C(10)-H(28)	54.9	67.3(26)	67.4(27)	68.1(27)	68.0(25)
p_{39}	ϕP(3)-C(8)-C(11)-H(31)	66.2	175.3(27)	175.3(27)	175.4(27)	175.4(25)
p_{40}	ϕP(3)-C(12)-C(13)-H(34)	-71.1	50.4(27)	50.0(27)	50.1(27)	50.6(25)
p_{41}	ϕP(3)-C(12)-C(14)-H(37)	-51.9	53.3(26)	52.5(27)	52.8(27)	51.9(25)
p_{42}	ϕP(3)-C(12)-C(15)-H(40)	-57.3	177.4(26)	176.3(27)	177.1(27)	176.2(25)

3.3 Results

3.3.1. Tri-*tert*-butylphosphine oxide

3.3.1.1. Theoretical methods

The structure of $OPBu^t_3$ was determined *ab initio*. A previous potential-energy surface search[5] involving rotation of the *tert*-butyl groups around the P-C bonds found only one conformer of $OPBu^t_3$, exhibiting C_3 symmetry. GED refinements can be complicated by the presence of multiple conformers, which we therefore wish to avoid at this stage. The presence of only one conformer thus makes this molecule well suited to be a test case for the new method. The molecular geometry of $OPBu^t_3$ at the MP2/6-311G* level can be found in Table 3.1.

At the outset of the SEMTEX refinement process, the heavy-atom positions were fixed as calculated using DYNAMITE, and both MP2/6-311G* and MM3 calculations were performed on the light atoms. The resulting light-atom parameters can be found in Table 3.3. As the values in Table 3.3 show, the C-C-H angles at the MP2 level of theory cover a range of 6.1°, while for the MM3 calculation this variation is far lower, at around 2.7°. Clearly, this is a significant difference in the structure as determined by these different theoretical methods. The inclusion of the *ab initio* data in the refinement *via* the SEMTEX method will allow this and other structural features to be modelled more accurately.

3.3.1.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311G* level. Fifteen geometric parameters were refined along with twenty-one groups of vibrational amplitudes. Five geometric and nine amplitude restraints were applied according to the SARACEN method. These can be found in Table 3.1. The final *R* factors for the refinement were found to be R_G = 0.061 and R_D = 0.089. Interatomic distances and corresponding amplitudes of vibration are given in Table S3.1 in the supplementary information, and final

experimental coordinates from the SARACEN GED analysis are given in Table S3.2. The least-squares correlation matrix is given in Table S3.3.

Table 3.3. C-H bond lengths, C-C-H bond angles and P-C-C-H bond torsions for $OPBu^t_3$ calculated with the MM3 and MP2/6-311G* methods. All bond lengths are in pm and angles in °. Internuclear distances are the calculated (r_e) values.

	*r*C-H		∠C-C-H		ϕP-C-C-H	
Parameter	MM3	MP2	MM3	MP2	MM3	MP2
P(2)-C(3)-C(4)-H(15)	111.3	109.6	111.1	108.2	-173.6	-168.5
P(2)-C(3)-C(4)-H(16)	110.7	108.9	113.1	113.0	67.3	72.5
P(2)-C(3)-C(4)-H(17)	111.1	109.1	112.3	110.4	-54.5	-49.3
P(2)-C(3)-C(5)-H(18)	111.3	109.5	111.9	108.4	173.8	175.7
P(2)-C(3)-C(5)-H(19)	111.1	109.3	111.7	111.5	53.7	57.5
P(2)-C(3)-C(5)-H(20)	110.7	109.1	113.0	114.0	-66.7	-65.1
P(2)-C(3)-C(6)-H(21)	111.3	109.5	111.5	107.9	173.8	173.4
P(2)-C(3)-C(6)-H(22)	111.1	109.3	111.1	111.9	56.0	55.8
P(2)-C(3)-C(6)-H(23)	110.6	108.8	113.8	113.1	-66.4	-67.7
range	0.7	0.8	2.7	6.1	N/A	N/A

3.3.1.3 DYNAMITE refinement

The starting parameters and force field were as for the SARACEN refinement, and all geometric parameters were refined according to this method. Once this was complete, the DYNAMITE code was activated and the light-atom positions were updated computationally. Consequently, the parameters associated with the hydrogen

atoms now represent average values over all atoms in a *tert*-butyl group. As for the SARACEN refinement, fifteen geometric parameters and twenty-one groups of vibrational amplitudes were refined. The final R factors for the refinement were found to be $R_G = 0.061$ and $R_D = 0.087$. Interatomic distances and corresponding amplitudes of vibration can be found in Table S3.4, and final experimental coordinates from the DYNAMITE GED analysis are given in Table S3.5. The least-squares correlation matrix is given in Table S2.6.

3.3.1.3 SEMTEX refinement

The starting parameters were as for the SARACEN and DYNAMITE refinements. The geometric parameters were refined using first the SARACEN, then the DYNAMITE method. Once all fifteen geometric parameters and twenty-one groups of vibrational amplitudes were refined according to the DYNAMITE method, the SEMTEX code was activated. The heavy-atom positions were fixed and theoretical structures calculated at both the MP2 and MM3 levels of theory. The differences in the light-atom parameters between these two structures were then calculated. During each refinement cycle for each parameter, the light-atom positions returned by the MM3 code were immediately modified by this set of differences.

As for the SARACEN and DYNAMITE refinements, all fifteen geometric parameters were refined along with twenty-one groups of vibrational amplitudes. Five geometric and nine amplitude restraints were applied using the SARACEN method. In the final refinement the R factors were $R_G = 0.062$ and $R_D = 0.086$. Figure 2.3 shows the radial distribution curve, and the molecular intensity curves are shown in Figure 3.4. Table 3.1 lists the final refined parameters, with the complete sets of light-atom parameters from DYNAMITE and SEMTEX listed in Table 3.4. Interatomic distances and corresponding amplitudes of vibration are given in Table S3.7 and final experimental coordinates from the SEMTEX GED analysis are given in Table S3.8. The least-squares correlation matrix is given in Table S3.9.

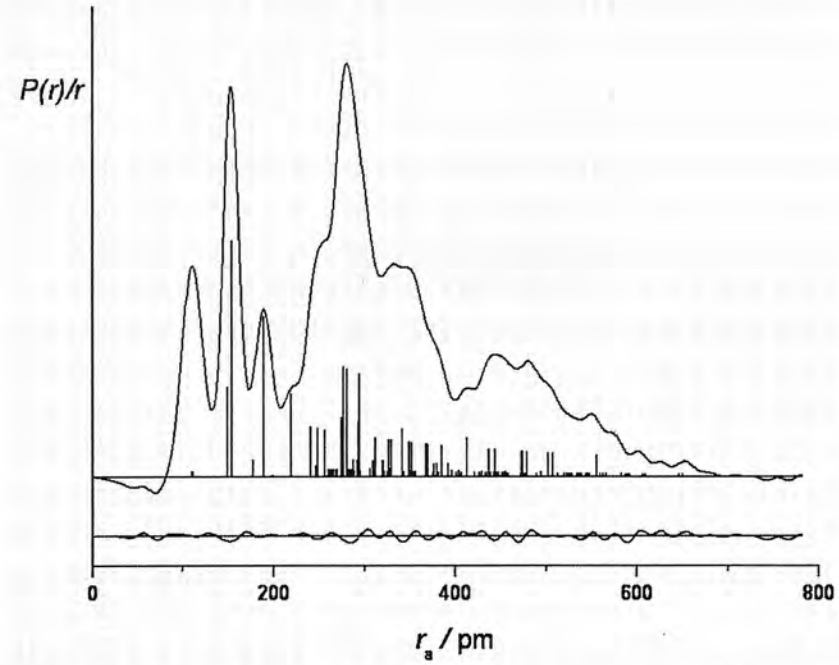


Figure 3.3. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $OPBu^t_3$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_P - f_P)$.

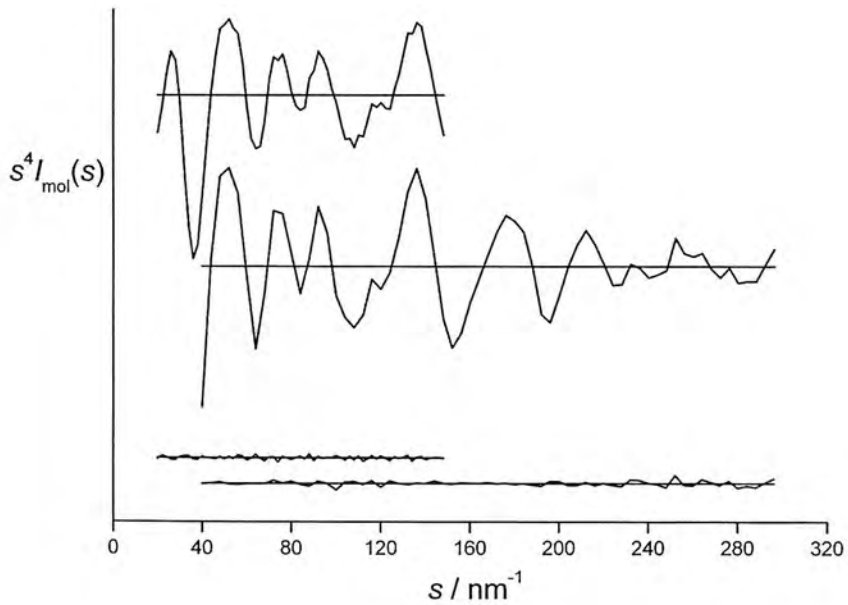


Figure 3.4. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $OPBu^t_3$.

Table 3.4. C-H bond lengths, C-C-H bond angles and P-C-C-H bond torsions for $OPBu^t_3$ for both the DYNAMITE (DYN) and SEMTEX (SEM) refinements (distances in pm, angles in °).

Parameter	rC-H (r_{h1})		∠C-C-H		ϕP-C-C-H	
	DYN	SEM	DYN	SEM	DYN	SEM
P(2)-C(3)-C(4)-H(15)	108.2	108.0	110.6	109.6	-167.5	-167.6
P(2)-C(3)-C(4)-H(16)	107.6	107.2	112.5	114.2	73.5	73.4
P(2)-C(3)-C(4)-H(17)	108.1	107.5	111.8	111.7	-48.3	-48.6
P(2)-C(3)-C(5)-H(18)	108.2	107.9	111.5	109.9	174.8	-178.8
P(2)-C(3)-C(5)-H(19)	108.1	107.7	112.4	112.8	54.6	61.4
P(2)-C(3)-C(5)-H(20)	107.6	107.5	111.1	115.4	-65.7	-59.0
P(2)-C(3)-C(6)-H(21)	108.2	107.9	111.1	109.3	168.6	169.3
P(2)-C(3)-C(6)-H(22)	108.1	107.8	110.6	113.2	50.0	51.3
P(2)-C(3)-C(6)-H(23)	107.5	107.2	113.3	114.4	-71.4	-71.3
Range	0.7	0.8	2.7	6.1	N/A	N/A

3.3.2. Tri-*tert*-butylphosphine imide

3.3.2.1. Theoretical methods

The structure of $HNPBu^t_3$ was determined using both *ab initio* and molecular mechanics methods. In this case, only one conformer, exhibiting C_1 symmetry, was found. As a consequence of the low symmetry of this molecule, the SEMTEX refinement process was considerably more time-consuming than for the $OPBu^t_3$ investigation. The molecular geometry of $HNPBu^t_3$ at the MP2/6-311G* level can be found in Table 3.2. The values of the C-H bond lengths, C-C-H bond angles and P-C-C-H bond torsions from the MP2/6-311G* calculation can be found in Table 3.5. As the values in Table 3.5 show, the C-C-H angles at the MP2 level of theory show a variation of 6.2°, while for the MM3 calculation this variation is 5.8°. This discrepancy between the MP2 and MM3 level calculations is less pronounced than for the $OPBu^t_3$ case.

Table 3.5. C-H bond lengths, C-C-H bond angles and P-C-C-H bond torsions for $HNPBu^t_3$ calculated with the MM3 and MP2/6-311G* methods. All bond lengths are in pm and angles in °. Interatomic distances are the calculated (r_e) values.

Parameter	*r*C-H		∠C-C-H		ϕP-C-C-H	
	MM3	MP2	MM3	MP2	MM3	MP2
P(3)-C(4)-C(5)-H(16)	110.8	108.8	112.8	113.7	76.1	74.5
P(3)-C(4)-C(5)-H(17)	111.2	108.9	113.4	111.2	-49.9	-50.7
P(3)-C(4)-C(5)-H(18)	111.3	109.6	110.8	107.7	-168.0	-168.5
P(3)-C(4)-C(6)-H(19)	110.7	108.9	113.4	113.8	-63.8	-64.2
P(3)-C(4)-C(6)-H(20)	111.3	109.1	112.4	110.0	57.5	57.1
P(3)-C(4)-C(6)-H(21)	111.3	109.5	111.8	107.7	177.2	176.5
P(3)-C(4)-C(7)-H(22)	111.3	109.8	110.7	109.1	-171.6	-171.4

Parameter	*r*C-H		∠C-C-H		*ϕ*P-C-C-H	
	MM3	MP2	MM3	MP2	MM3	MP2
P(3)-C(4)-C(7)-H(23)	110.8	109.1	113.0	111.8	70.3	70.8
P(3)-C(4)-C(7)-H(24)	110.6	108.8	114.3	113.9	-52.2	-51.9
P(3)-C(8)-C(9)-H(25)	110.8	109.0	113.1	113.2	72.4	73.3
P(3)-C(8)-C(9)-H(26)	111.3	109.6	111.3	108.0	-169.4	-168.1
P(3)-C(8)-C(9)-H(27)	111.2	109.3	112.5	111.1	-50.7	-49.8
P(3)-C(8)-C(10)-H(28)	111.2	109.3	112.6	110.9	55.3	55.2
P(3)-C(8)-C(10)-H(29)	111.3	109.5	111.6	108.3	174.6	173.4
P(3)-C(8)-C(10)-H(30)	110.8	109.0	113.3	113.6	-66.4	-67.4
P(3)-C(8)-C(11)-H(31)	111.1	109.4	112.0	112.0	61.9	62.5
P(3)-C(8)-C(11)-H(32)	111.3	109.6	111.8	108.9	-178.6	-178.9
P(3)-C(8)-C(11)-H(33)	111.0	109.3	112.4	112.7	-58.7	-59.7
P(3)-C(12)-C(13)-H(34)	110.5	108.3	114.9	113.6	-87.3	-89.9
P(3)-C(12)-C(13)-H(35)	111.3	109.6	109.4	107.8	157.3	152.9
P(3)-C(12)-C(13)-H(36)	111.1	109.9	115.2	112.2	39.9	34.7
P(3)-C(12)-C(14)-H(37)	111.2	109.2	113.9	110.6	-49.2	-49.9
P(3)-C(12)-C(14)-H(38)	111.3	109.7	110.7	109.1	-169.0	-168.0
P(3)-C(12)-C(14)-H(39)	111.0	109.2	111.7	113.0	74.1	72.8
P(3)-C(12)-C(15)-H(40)	111.0	109.3	112.4	113.1	-56.1	-57.3
P(3)-C(12)-C(15)-H(41)	111.3	109.9	110.0	109.5	-174.0	-177.0

Parameter	*r*C-H		∠C-C-H		*ϕ*P-C-C-H	
	MM3	MP2	MM3	MP2	MM3	MP2
P(3)-C(12)-C(15)-H(42)	110.8	108.7	113.9	110.7	67.6	65.2
range	0.8	1.6	5.8	6.2	N/A	N/A

3.3.2.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311G* level. In total forty-two geometric parameters were refined along with ten groups of vibrational amplitudes. Thirty-five geometric and six amplitude restraints were applied according to the SARACEN method. These can be found in Table S3.10. The final *R* factors for the refinement were found to be $R_G = 0.072$ ($R_D = 0.097$). Final refined parameters can be found in Table 3.2. Interatomic distances and corresponding amplitudes of vibration are given in Table S3.10, with final experimental coordinates from the SARACEN GED analysis given in Table S3.11. The least-squares correlation matrix is given in Table S3.12.

3.3.2.3. DYNAMITE refinement

The starting parameters and force field were as for the SARACEN refinement, and all geometric parameters were refined according to this method. Once this was complete, the DYNAMITE code was activated and the light-atom positions were updated computationally. Consequently, the parameters associated with the hydrogen atoms now represent average values over all nine atoms in a *tert*-butyl group. As for the SARACEN refinement, forty-two geometric parameters and ten groups of vibrational amplitudes were refined. The final *R* factors for the refinement were found to be $R_G = 0.068$ ($R_D = 0.097$). Final refined parameters are listed in Table 3.2. Interatomic distances and corresponding amplitudes of vibration can be found in Table S3.13, with the final experimental coordinates from the DYNAMITE GED

analysis given in Table S3.14. The least-squares correlation matrix is given in Table S3.15.

3.3.2.4. SEMTEX refinement

The starting parameters were as for the SARACEN and DYNAMITE refinements. The geometric parameters were refined using first the SARACEN, then the DYNAMITE method. As for the $OPBu^t_3$ case, the SEMTEX code was then activated. As for the SARACEN and DYNAMITE refinements, all forty-two geometric parameters were refined along with ten groups of vibrational amplitudes. Five geometric and nine amplitude restraints were applied using the SARACEN method. In the final refinement the *R* factors were $R_G = 0.068$ ($R_D = 0.097$). Figure 3.5 shows the final radial distribution curve from the refinement, and the molecular intensity curves are shown in Figure 3.6. Table 3.2 lists the final refined parameters. The light-atom parameters determined using DYNAMITE and SEMTEX are given in Table 3.6. Interatomic distances and corresponding amplitudes of vibration are given in Table S3.16, and final experimental coordinates from the SEMTEX refinement are given in Table S3.17. The least-squares correlation matrix is given in Table S3.18.

Table 3.6. C-H bond lengths, C-C-H bond angles and P-C-C-H bond torsions for $HNPBu^t_3$ for both the DYNAMITE (DYN) and SEMTEX (SEM) refinements (distances in pm, angles in °).

Parameter	*r*C-H (r_{h1})		∠C-C-H		ϕP-C-C-H	
	DYN	SEM	DYN	SEM	DYN	SEM
P(3)-C(4)-C(5)-H(16)	114.6	114.6	110.8	113.2	74.7	74.8
P(3)-C(4)-C(5)-H(17)	115.0	114.7	111.3	110.7	-51.1	-51.2
P(3)-C(4)-C(5)-H(18)	115.2	115.4	108.9	107.0	-169.3	-169.3
P(3)-C(4)-C(6)-H(19)	114.6	114.7	111.3	113.2	-70.2	-70.5

Parameter	rC-H (r_{h1})		∠C-C-H		ϕP-C-C-H	
	DYN	SEM	DYN	SEM	DYN	SEM
P(3)-C(4)-C(6)-H(20)	115.1	114.9	110.4	109.5	51.1	50.8
P(3)-C(4)-C(6)-H(21)	115.2	115.3	109.8	107.3	170.8	170.5
P(3)-C(4)-C(7)-H(22)	115.2	115.6	108.8	108.7	-176.7	-176.9
P(3)-C(4)-C(7)-H(23)	114.7	114.9	111.0	111.3	65.2	64.9
P(3)-C(4)-C(7)-H(24)	114.5	114.6	112.2	113.3	-57.3	-57.6
P(3)-C(8)-C(9)-H(25)	114.7	114.8	111.1	112.7	-177.8	-178.0
P(3)-C(8)-C(9)-H(26)	115.2	115.4	109.4	107.5	-59.7	-59.8
P(3)-C(8)-C(9)-H(27)	115.1	115.0	110.5	110.6	59.1	58.9
P(3)-C(8)-C(10)-H(28)	115.1	115.1	110.6	110.5	83.1	83.0
P(3)-C(8)-C(10)-H(29)	115.2	115.3	109.6	107.9	-157.6	-157.7
P(3)-C(8)-C(10)-H(30)	114.7	114.8	111.3	113.0	-38.5	-38.6
P(3)-C(8)-C(11)-H(31)	115.0	115.2	110.0	111.5	-29.6	-29.4
P(3)-C(8)-C(11)-H(32)	115.1	115.4	109.8	108.4	89.9	90.1
P(3)-C(8)-C(11)-H(33)	114.9	115.1	110.4	112.1	-150.2	-150.1
P(3)-C(12)-C(13)-H(34)	114.3	114.1	112.8	113.1	-165.9	-165.8
P(3)-C(12)-C(13)-H(35)	115.2	115.7	107.5	107.3	78.7	78.9
P(3)-C(12)-C(13)-H(36)	114.9	114.8	113.2	111.7	-38.9	-38.6
P(3)-C(12)-C(14)-H(37)	115.1	115.0	111.8	110.1	60.1	58.9
P(3)-C(12)-C(14)-H(38)	115.1	115.5	108.8	108.6	-59.7	-60.8

Parameter	rC-H (r_{h1})		∠C-C-H		ϕP-C-C-H	
	DYN	SEM	DYN	SEM	DYN	SEM
P(3)-C(12)-C(14)-H(39)	114.9	115.0	109.6	112.5	-176.6	-177.7
P(3)-C(12)-C(15)-H(40)	114.9	115.0	110.4	112.6	-59.9	-60.4
P(3)-C(12)-C(15)-H(41)	115.1	115.7	108.1	109.1	-177.8	-178.3
P(3)-C(12)-C(15)-H(42)	114.7	114.4	111.8	110.3	63.7	63.2
Range	0.9	1.6	5.7	6.3	N/A	N/A

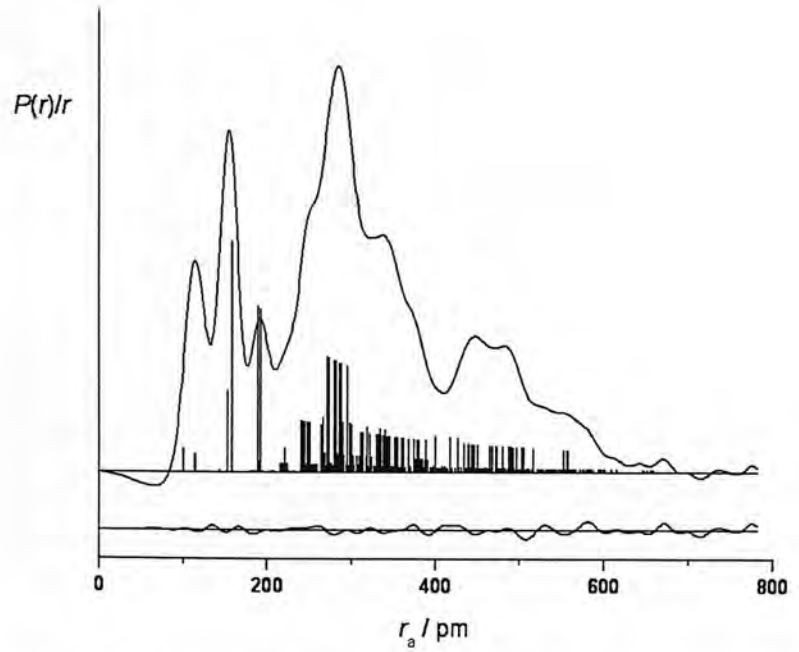


Figure 3.5. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $HNPBu^t_3$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_P - f_P)$.

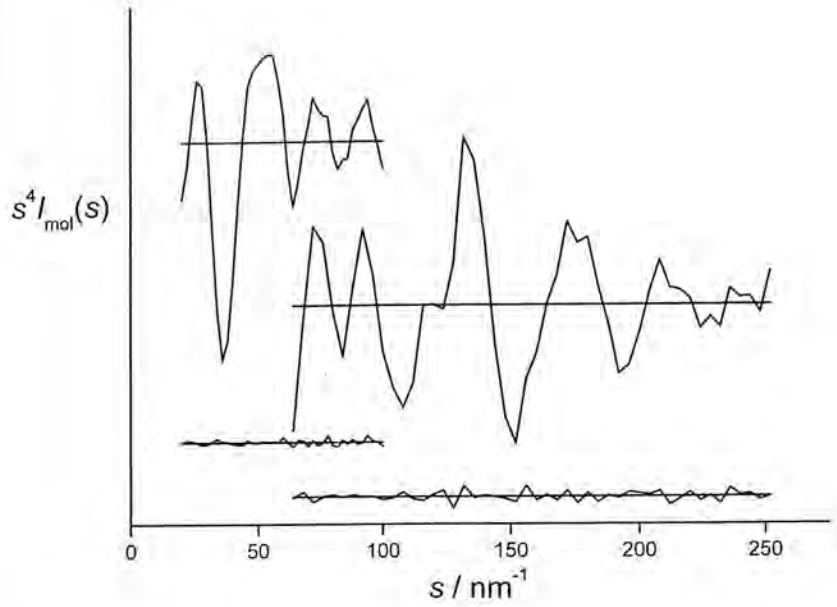


Figure 3.6. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $OPBu^t_3$.

3.4. Discussion

The molecular structures of tri-*tert*-butylphosphine oxide and tri-*tert*-butylphosphine imide were re-examined as test cases for the new SEMTEX method of gas electron diffraction structure refinement. Previously, these molecules had been studied using the DYNAMITE total structure determination method, which uses molecular mechanics to model the positions of the hydrogen atoms throughout the refinement process.

The new SEMTEX method goes one step further by using high-level theoretical data, in this case at the MP2/6-311G* level, dynamically within the refinement process. In the case of $OPBu^t_3$, large discrepancies were observed between the molecular mechanics and *ab initio* calculated structures. A particularly notable example of this was the range of CCH angles, which was more than twice as large for the MP2 case as for MM3: 6.1° as compared to 2.7°.

For the SEMTEX refinement of $OPBu^t_3$, experimental and theoretical parameters are generally in good agreement with each other. The C-C bonded distance refined to a value of 154.0 pm, compared with 153.8 pm calculated at the MP2/6-311G* level of theory. Angles also generally agreed to within 1.5°. For example, ∠O-P-C refined to 107.7° compared to 109.1° from calculations.

There is a very good level of agreement between the DYNAMITE and SEMTEX refinements, as was expected in this case due to the relatively simple nature of the molecule under investigation. For the heavy atoms, all bonded distance parameters agree to within 0.5 pm. The largest heavy-atom discrepancy occurs in the P-C bonded distance, a difference of only 0.3 pm. Angles also are in close agreement between the two methods.

The average light-atom parameters (C-C distance, C-C-H angle and P-C-C-H torsion averaged over nine hydrogen atoms in a *tert*-butyl group) also agree well. However, a notable difference between the structures of the DYNAMITE and SEMTEX refinements is the range of different values for these parameters. For the DYNAMITE refinement, the range of C-C-H angles was found to be 2.7°, while for the SEMTEX refinement, it was more than double this, at 6.1°. This reflects the difference between the *ab initio* and molecular mechanics structures mentioned above, and shows that the MM3 method used previously did not allow for the complete asymmetry that the structure should adopt.

For $HNPBu^t_3$, the range of CCH angles was found to be 6.2° for the MP2 calculation, and 5.8° for the MM3. This is clearly much less of a difference in ranges when compared to the $OPBu^t_3$ structure. In this case, less difference would therefore be expected in the structures of the DYNAMITE and SEMTEX refinements. This prediction is borne out in the final refined structures, where a range of 6.3° is found for the SEMTEX refinement, in comparison with a 5.7° range for DYNAMITE.

As in the $OPBu^t_3$ case, there is very close agreement of parameters given by the DYNAMITE and SEMTEX methods. The bond lengths all agree to within 0.3 pm, with the average C-H distance increased by this amount. Experimental and theoretical parameters are also in very good agreement. For example, the average N-P-C angle is determined to be 106.5° from *ab initio*, and 107.2° by SEMTEX. The

light-atom parameters also agree well, with, for example, 0.3 pm variation in the average C-H distance between DYNAMITE and SEMTEX.

3.5. Conclusion

The structures of two sterically-crowded molecules, $OPBu^t_3$ and $HNPBu^t_3$, were re-determined as test cases for the new SEMTEX method of total structure determination. As expected, the fit to experimental data changed little between the DYNAMITE and SEMTEX refinements for these relatively simple cases. However, the structures were finally able to adopt the completely asymmetric conformations that are shown by *ab initio* theoretical methods to be desirable. The success of the refinements of these two structures demonstrated that the SEMTEX method was functioning correctly, and could subsequently be applied to more challenging and complex molecules where it could be expected to make a bigger difference to the overall quality of refinements.

3.6. References

1. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Dalton Trans.* 2004, 384.

2. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Inorg. Chem.*, 2004, **43**, 5522.

3. D. W. H. Rankin, H. E. Robertson, R. Seip, H. Schmidbaur and G. Blaschke, *J. Chem. Soc., Dalton Trans.* 1985, 827.

4. A. W. Ross, M. Fink and R. Hilderbrandt, *International Tables for Crystallography*, ed. A. J. C. Wilson, Kluwer Academic Publishers, Dordrecht, Boston and London, 1992, **vol. C**, p. 245.

4. Comparison of the molecular structure of perfluoromethylcyclohexane ($C_6F_{11}CF_3$) as determined using the SARACEN, DYNAMITE and SEMTEX methods

4.1. Introduction

The successful testing of the SEMTEX code on $OPBu^t_3$ and $HNPBu^t_3$ illustrated that the technique was build on a sound methodology. At this stage however, the method had only been tested with hydrogen atoms in relatively simple cases, where it had no great effect on the overall goodness of fit. It also (like the DYNAMITE method at this stage of development) was only applicable to peripheral atoms in a branched chain structural environment.

Consequently, the next challenges were twofold. First, it was desirable to expand the scope of the method, to be able to treat peripheral atoms in an environment other than a branched chain. The next stage was to identify and attempt to refine the structure of a molecule to which it could conceivably make a substantial improvement over the DYNAMITE method.

With these two considerations in mind, the perfluorinated ring alkane molecule perfluoromethylcyclohexane ($C_6F_{11}CF_3$) was determined to be a good candidate for study. Such molecules are known to be highly chemically inert, since they contain strong C–F bonds,[1] and as such are potential contributors to the destruction of the ozone layer. They are also very hydrophobic and thermally stable, and have been proposed as candidates for a wide range of uses, from oxygen carriers[2] to surfactants.[3]

In this work, the SEMTEX technique has been further developed to allow it to also be applied to ring structures. This chapter presents the molecular structure of $C_6F_{11}CF_3$ as determined in the gas phase using the new method. The original DYNAMITE[4] code has also been been expanded upon in order to deal with peripheral atoms in a ring environment.

4.2. Experimental

Two conformers of $C_6F_{11}CF_3$ were located, with the methyl group in the equatorial position in one conformer and the axial position in the other. Both conformers were found to have C_S symmetry. The difference in energy between the conformers was calculated to be sufficiently high that only one conformer could be expected to be

present in significant quantity in the experimental gas-phase sample mixture – that with the trifluoromethyl group in the equatorial position. Therefore, all subsequent geometry optimisations were carried out only on this conformer. The amplitudes of vibration (u_{h1}) and curvilinear vibrational correction factors (k_{h1}) to distances were derived as detailed in Chapter 2. The lowest-energy structure is shown in Figure 4.1.

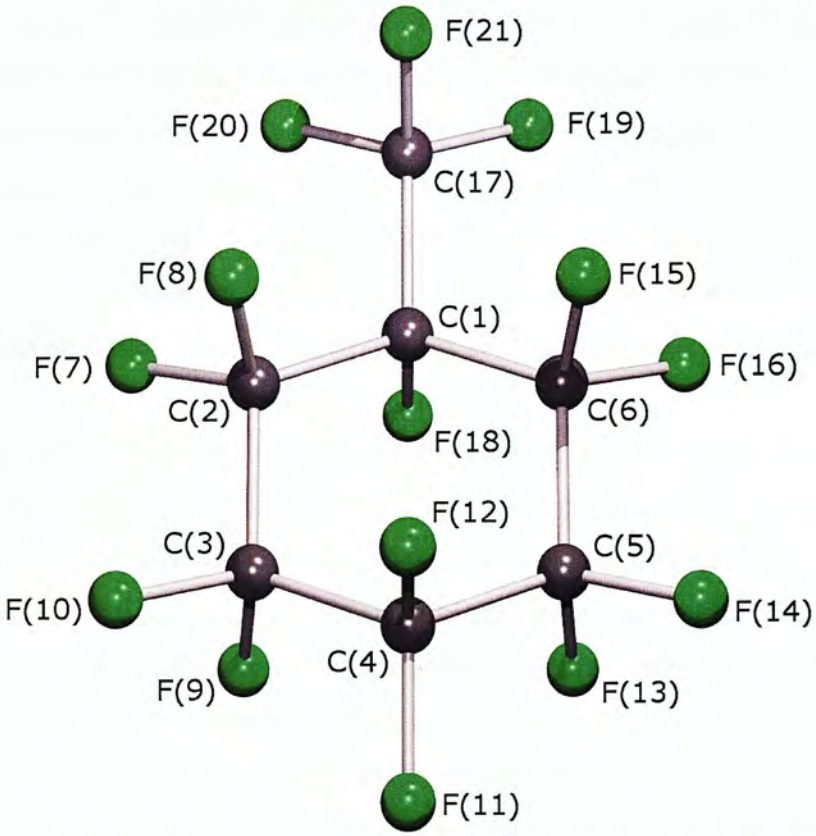


Figure 4.1 – Gas-phase molecular structure of $C_6F_{11}CF_3$ viewed with the *C*s plane of symmetry vertical.

The electron diffraction data was collected as detailed in Chapter 2. The weighting points for the off-diagonal weight matrices, correlation parameters and scale factors for both distances are given in the Supplementary information.

4.2.1. Electron diffraction model

The structure was defined using a model with C_S symmetry, as indicated by the *ab initio* calculations. Seven independent geometric parameters were required, comprising two bond lengths, three bond angles and differences and two torsion

parameters and differences. These can be found in Table 4.1. The starting values used for the parameters in the model are the r_e values obtained from the MP2/6-311G* calculation.

The core-atom bond lengths were described by *r*C-C av. (p_1). *Ab initio* calculations showed that in principle there are many different C-C distances in the structure. However, the differences are small and largely invariant on refinement, and were therefore fixed at values computed at the MP2/6-311+G* level in the electron diffraction model to ensure a workable number of refining parameters.

For the core-atom bond angles, the average of all C-C-C angles, ∠C-C-C av. (p_4) was used, again with the small differences between them fixed to values calculated *ab initio*. Two parameters were also used to describe the chair conformation of the C_6-ring, the average dihedral "drop" angle describing the displacement of C_1 and C_4 from the plane containing C_2, C_3, C_5 and C_6, and the difference between these two angles.

Three parameters were necessary to describe the starting positions of the fluorine atoms. These comprised the average *r*C-F (p_5), and an average and difference of C-C-F angles ($p_{6\text{-}7}$). These parameters represent the mean value for all eight F atoms in one half of a $C_6F_{11}CF_3$ molecule. For the SARACEN refinement, the fluorine atoms are restricted to these average parameter values, modified where necessary by small fixed differences from the initial MP2 calculation, since to include each distance and pair of angles explicitly in the refinement would lead to an unworkable number of refining parameters. In the DYNAMITE and SEMTEX refinements, however, the absolute values for each individual atom are determined using the respective computational methods (MM3 and MP2), and then scaled to the values of these average parameters in the ongoing GED refinement.

4.2.2. SEMTEX method for peripheral atoms in a ring structure

In order to apply DYNAMITE and SEMTEX to the fluorine atoms in $C_6F_{11}CF_3$, it was necessary to expand the method to handle peripheral atoms in a ring environment. A different set of parameters is needed to describe the atoms in this case than the set used for the hydrogen atoms in the previously-discussed phosphine

molecules. Instead of a bond length, bond angle and torsional angle, the individual ring F-atom positions in $C_6F_{11}CF_3$ are described using their F-C bond length, and two F-C-C angles, one in each direction around the ring. These were then scaled back to the average parameter values from the ongoing GED refinement, as for the branched-chain cases previously studied. In this case, these parameters are *r*C-F (p_2) for the distances, and ∠C-C-F av. and ∠C-C-F diff. (p_{4-5}) for the angles. The F atoms in the CF_3 group are still in a branched-chain local environment, and could thus be treated in the same way as the H atoms in the phosphine refinements.

4.3. Results

4.3.1. Theoretical methods

The structure of $C_6F_{11}CF_3$ was determined *ab initio*. Initial calculations suggested two possible conformers, with the CF_3 group in either the axial or equatorial position. The equatorial conformer was calculated to be 24.7 kJ mol^{-1} lower in energy than the axial at the HF/3-21G* level, however, and it was therefore decided that only this conformer would be present in the experimental mixture in significant quantity. The molecular geometry of $C_6F_{11}CF_3$ at the MP2/6-311+G* level can be found in Table 4.1.

At the outset of the SEMTEX refinement process, the heavy-atom positions were fixed as calculated using DYNAMITE, and both MP2/6-311+G* and MM3 calculations were performed on the light atoms. The resulting light-atom parameters can be found in Table 4.2.

4.3.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311+G* level. All seven independent geometric parameters were refined, along with thirteen groups of vibrational amplitudes. One geometric and eight amplitude constraints were applied according to the SARACEN method. The geometric parameters and restraints can be found in Table 4.1. The final

R factors for the refinement were found to be $R_G = 0.063$ and $R_D = 0.040$. Interatomic distances and corresponding amplitudes of vibration are given in the supplementary information along with final experimental coordinates from the SARACEN GED analysis.

4.3.3. DYNAMITE refinement

The starting parameters and force field were as for the SARACEN refinement, and all geometric parameters were refined according to this method. Once this was complete, the DYNAMITE code was activated and the light-atom positions were updated computationally. Consequently, the parameters associated with the fluorine atoms now represent average values over all such atoms in the structure.

As for the SARACEN refinement, seven geometric parameters were refined, along with thirteen groups of vibrational amplitudes. One geometric and eight amplitude restraints were applied. The final R factors for the refinement were found to be $R_G = 0.133$ and $R_D = 0.076$. The refined GED parameters can be found in Table 4.1. Interatomic distances and corresponding amplitudes of vibration can be found in the supplementary information, along with final experimental coordinates from the DYNAMITE GED analysis.

4.3.4. SEMTEX refinement

The starting parameters were as for the SARACEN and DYNAMITE refinements. The geometric parameters were refined using first the SARACEN, then the DYNAMITE method. Once all seven geometric parameters and thirteen groups of vibrational amplitudes were refined according to the DYNAMITE method, the SEMTEX code was activated. The heavy-atom positions were fixed and theoretical structures calculated at both the MP2 and MM3 levels of theory. The differences in the peripheral-atom parameters between these two structures were then calculated. During each refinement cycle for each parameter, the peripheral-atom positions returned by the MM3 code were immediately modified by this set of differences.

As for the SARACEN and DYNAMITE refinements, all seven geometric parameters were refined, along with thirteen groups of vibrational amplitudes. One geometric and eight amplitude restraints were again applied using the SARACEN method. In the final refinement the R factors were $R_G = 0.062$ and $R_D = 0.038$. The molecular intensity curve for $C_6F_{11}CF_3$ is given in Figure 4.2, and Figure 4.3 shows the radial distribution curve for the refinement. The refined geometric parameters are given in Table 4.1. Table 4.2 gives the light-atom parameters determined using both MM3 and MP2, while Table 4.3 gives the final light-atom parameters from both DYNAMITE and SEMTEX. Interatomic distances and corresponding amplitudes of vibration are given in the supplementary information, along with final experimental coordinates from the SEMTEX GED analysis.

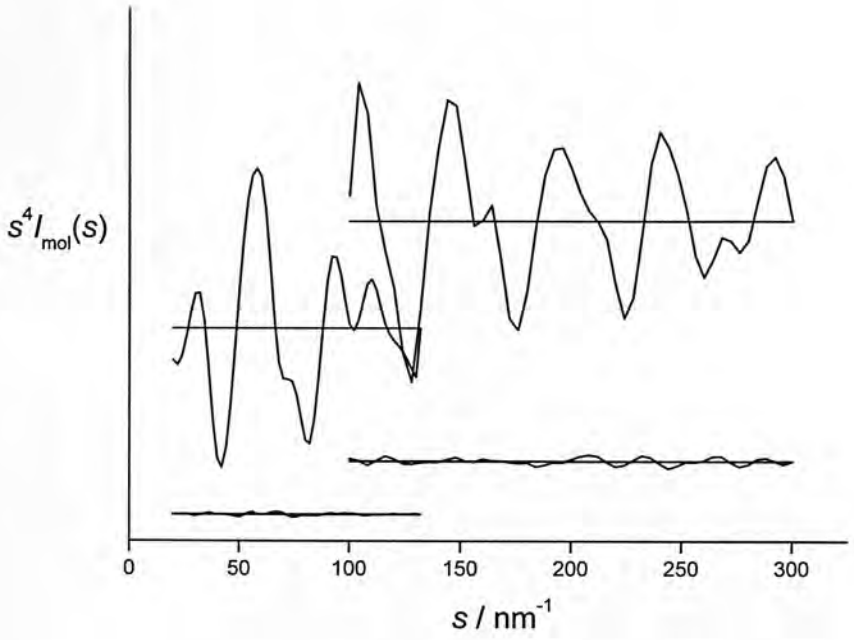


Figure 4.2. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $C_6F_{11}CF_3$.

Table 4.1 – GED refined parameters for the the SARACEN, DYNAMITE and SEMTEX refinements of $C_6F_{11}CF_3$. All bond lengths are in pm and angles in °.

No.	Parameter	MP2/ 6-311+G*(r_e)	SARACEN(r_{h1})	DYNAMITE(r_{h1})	SEMTEX(r_{h1})	Restraint
p_1	*r*C-C	154.5	154.9(2)	154.3(2)	154.8(2)	
p_2	*r*C-F	134.0	133.8(1)	133.9(1)	134.0(1)	
p_3	∠C-C-C av.	112.5	113.2(3)	110.4(4)	113.1(3)	
p_4	∠C-C-F av.	108.7	109.2(1)	109.4(1)	109.6(2)	
p_5	∠C-C-F diff.	1.2	-0.3(4)	1.2(5)	1.2(5)	
p_6	*dih.drop av.*	134.4	134.6(3)	139.3(5)	132.4(4)	
p_7	*dih.drop diff.*	1.2	1.2(5)	1.2(5)	1.0(5)	1.0(3)

Table 4.2 – C-F bond lengths and two sets of C-C-F bond angles for $C_6F_{11}CF_3$ calculated with the MM3 and MP2/6-311+G* methods. All bond lengths are in pm and angles in °. Internuclear distances are the calculated (r_e) values.

Parameter	*r*C(n)–F		∠C(n+1)–C(n)1–F		∠C(n-1)–C(n)–F	
[F-C-C(a)/C(b)]	MM	MP2	MM	MP2	MM	MP2
F(7)-C(2)-C(3)/C(1)	134.2	134.3	109.5	107.0	109.5	109.4
F(8)-C(2)-C(3)/C(1)	136.7	134.9	106.5	107.2	108.4	109.2
F(9)-C(3)-C(4)/C(2)	134.0	134.7	107.7	108.3	107.3	108.0
F(10)-C(3)-C(4)/C(2)	137.4	134.2	109.6	109.1	110.0	108.1
F(11)-C(4)-C(5)/C(3)	134.2	134.0	111.6	109.5	111.6	109.5
F(12)-C(4)-C(5)/C(3)	136.9	134.6	108.6	108.8	108.6	108.8
F(18)-C(1)-C(2)/C(6)	131.1	136.2	109.0	108.4	109.0	108.4

Parameter	*r*C-F		∠C-C-F (a)	
F(20)-C(17)-C(1)	136.0	133.3	110.7	108.9
F(21)-C(17)-C(1)	138.9	132.5	114.3	113.2

Table 4.3 - C-F bond lengths and two sets of C-C-F bond angles for $C_6F_{11}CF_3$ for both the DYNAMITE and SEMTEX refinements. All bond lengths are in pm and angles in °.

Parameter	rC(n)–F(r_{h1})		∠C(n+1)–C(n)–F		∠C(n-1)–C(n)–F	
[F-C-C(a)/C(b)]	DYNAMITE	SEMTEX	DYNAMITE	SEMTEX	DYNAMITE	SEMTEX
F(7)-C(2)-C(3)/C(1)	132.5	133.9	109.3	107.5	109.4	109.9
F(8)-C(2)-C(3)/C(1)	135.0	134.5	106.3	107.6	108.3	109.6
F(9)-C(3)-C(4)/C(2)	132.3	134.4	107.5	108.7	107.1	108.5
F(10)-C(3)-C(4)/C(2)	135.6	133.9	109.5	109.5	109.8	108.5
F(11)-C(4)-C(5)/C(3)	132.5	133.6	111.4	111.0	111.4	110.0
F(12)-C(4)-C(5)/C(3)	135.2	134.2	108.4	109.3	108.4	109.3
F(18)-C(1)-C(2)/C(6)	129.5	135.9	108.8	108.8	108.8	108.8
Parameter	rC-F(r_{h1})		∠C-C-F			
F(20)-C(17)-C(1)	134.3	133.0	110.3	109.4		
F(21)-C(17)-C(1)	137.1	132.2	114.1	113.7		

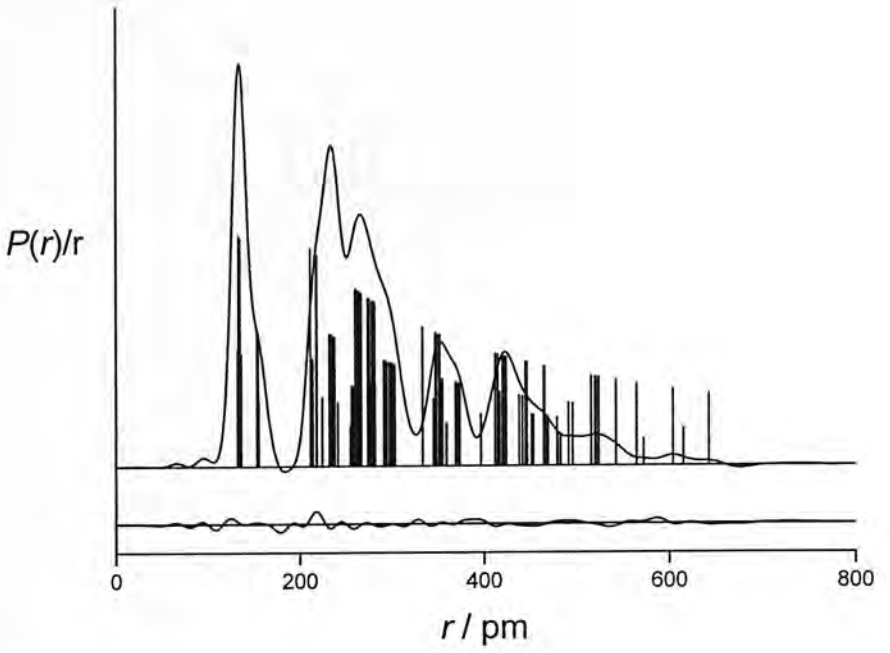


Figure 4.3. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $C_6F_{11}CF_3$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_P - f_P)$.

4.4 Discussion

The gas-phase structure of perfluoromethylcyclohexane has been examined using the SARACEN, DYNAMITE and SEMTEX methods of structure determination. The use of SEMTEX has allowed the full range of peripheral-atom parameters to be defined, free of constraints that were necessary in the SARACEN refinement of the structure.

For both the SARACEN and the SEMTEX refinements, refined parameters for the central atoms are generally in good agreement both with each other, and with the calculated MP2 structure. The C-C bonded distance refined to a value of 154.9(2) pm using SARACEN and 154.8(2) pm using SEMTEX, compared with 154.5 pm calculated at the MP2/6-311+G* level of theory. Angles also generally agreed to within 1.0°. For example, the average C-C-C angle refined to 113.2(3)° using

SARACEN and 113.1(3)° using SEMTEX , as compared with 112.6° from the calculated MP2 structure.

Two of the three average peripheral-atom parameters, the C-F distance and the average of C-C-F angles, are also in good agreement between the SARACEN and SEMTEX methods. The third however, ∠ C-C-F diff., differs considerably – -0.3(4)° for SARACEN compared with 1.2(5)° for SEMTEX. Comparison of these values with the value of 1.2° calculated at the MP2 level leads to the expected conclusion that the SEMTEX refinement better describes the positions of the peripheral atoms. The slight improvement in goodness-of-fit parameter exhibited by the SEMTEX refinement may reflect this difference.

The most notable result is the very poor agreement between the DYNAMITE refinement and both the SARACEN and SEMTEX refinements, and the high *R*-factor associated with the DYNAMITE case. Several parameters are not in good agreement between DYNAMITE and SEMTEX. For example, the average C-C distance refines to 154.3(3) pm using DYNAMITE, and 154.8(2) pm with SEMTEX. This disagreement between structures is most evident in the parameter ∠ C-C-C av., which refines to a similar value for SARACEN [113.2(3)°] and SEMTEX [113.1(3)°], but takes a completely different value in the DYNAMITE refinement, of 110.4(4)°. This is likely to be caused by the molecular mechanics calculating inaccurate positions for the peripheral F-atoms, and consequently these erroneously-placed peripheral atoms having a serious effect on the structure of the central C_6 ring. The dihedral drop angle is also substantially different between the DYNAMITE and SEMTEX cases (139.3(5)° and 132.4(4)° respectively) – this is also likely to be influenced by the positions of the peripheral F atoms.

The peripheral-atom parameters given in Table 4.3 give some indication of the difference between the MM and MP2 calculations in this case – for example, ∠ F(7)-C(2)-C(3) and ∠ F(8)-C(2)-C(3) differ in value by 3.0° in the MM case, and by only 0.1° in MP2. Furthermore, we should expect a geminal bond shortening effect to be evident in the individual values of the C-F distances. This is clearly apparent in the refined SEMTEX structure: the C-F distances in the CF_3 group are considerably shorter than those of the ring carbons, to which only two fluorine atoms are attached. The atom C(1), to which only one fluorine atom is attached, exhibits the longest C-F

distance of all. In contrast, the DYNAMITE structure does not show any evidence of this effect at all. This illustrates that the molecular mechanics method is unable to predict any geminal bond effect on the C-F distances, and is further evidence that it is insufficient to return accurate peripheral-atom positions in the refinement.

The difference in the final *R*-factors between the DYNAMITE and SEMTEX refinements is extremely marked, $R_G = 0.133$ for DYNAMITE and $R_G = 0.062$ for SEMTEX. This structure therefore illustrates very well the limitations of the molecular mechanics as used in DYNAMITE, and provides a first definite example of case where inclusion of *ab initio* light-atom data *via* the SEMTEX method can make a significant difference to the overall quality of refinement.

4.5 Conclusion

The structure of perfluoromethylcyclohexane, $C_6F_{11}CF_3$, has been determined using each of the SARACEN, DYNAMITE and SEMTEX methods and the results compared. The goodness-of-fit parameter R_G is similar between the SARACEN and SEMTEX refinements, with the SEMTEX method performing better in the prediction of the peripheral-atom locations. The most notable result is the relatively poor performance of the DYNAMITE method in determining this structure, producing markedly different carbon atom structure with an R_G-value greater than double that of the other two methods. The refinement of $C_6F_{11}CF_3$ therefore perfectly illustrates the need for the new SEMTEX method if accurate total structure determination of a wide range of molecules is desired.

4.6 References

1. A. R. Ravishankara, S. Solomon, A. A. Turnipseed and R. F. Warren, *Science,* 1993, **259**, 194.
2. J. G. Reiss, *Chem. Rev.*, 2001, **101**, 2797.
3. J. Eastoe, Z. Bayazit, S. Martel, D. C. Steytler and R. K. Heenan, *Langmuir*, 1996, **12**, 1423.
4. S. L. Hinchley, M. F. Haddow and D. W. H. Rankin, *Dalton Trans.*, 2004, 384.

5. The gas-phase structure of $W(NBu^t)_2(NHBu^t)_2$

5.1. Introduction

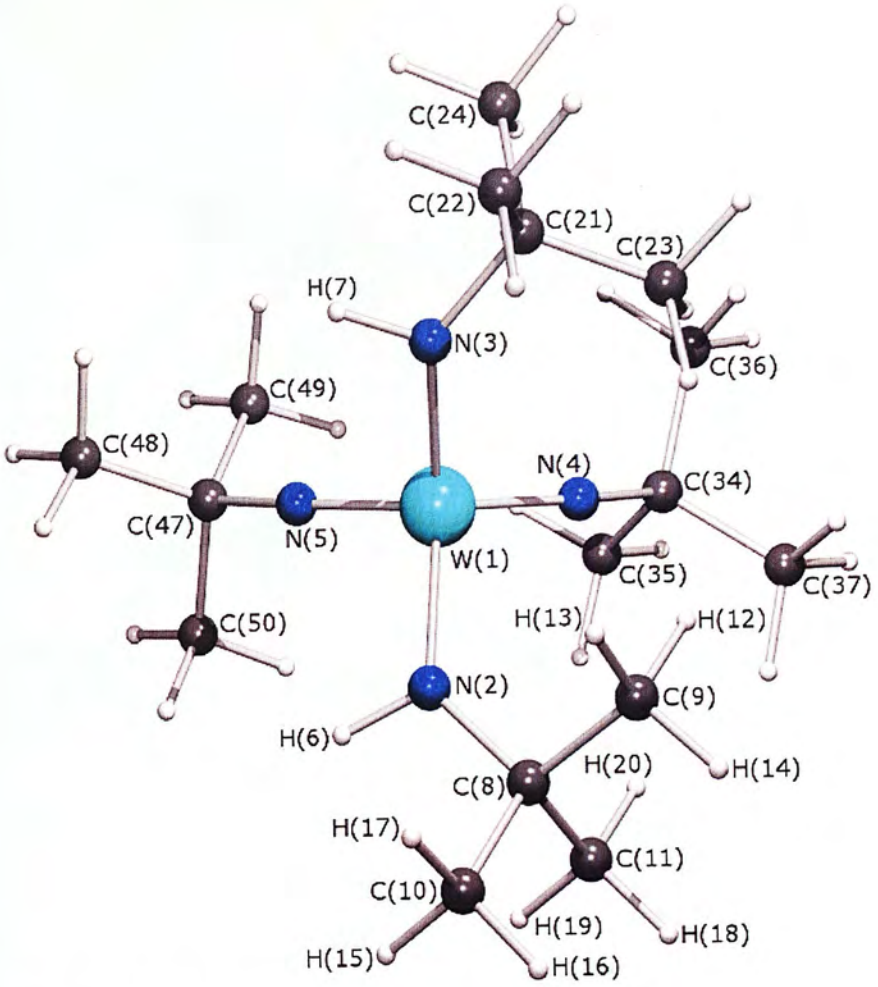
Tungsten imido complexes have played an important role in the development of the chemistry of W-N – containing compounds[1] and, in more recent years, have featured in materials chemistry research focussed on the formation of tungsten nitride, WN. The parent compounds in this class, the mixed imido(amido) species $W(NR)_2(NHR)_2$, were first described by Nugent (R = Bu^t),[2,3] though their chemistry owes much to the work of Wilkinson, who first isolated $Li_2[W(NBu^t)_4]$[4] and utilised this salt to prepare other hetero-bimetallic imido complexes.[5,6]

This work is concerned with determining the structure of $W(NBu^t)_2(NHBu^t)_2$ in the gas phase, which is of central importance to CVD studies.[7,8,9] The structure has been determined using each of the SARACEN, DYNAMITE and SEMTEX methods in order to ensure that the SEMTEX code was being properly applied to multi-conformer systems. It was not expected that DYNAMITE or SEMTEX would make any substantial difference to the refined structure in this case.

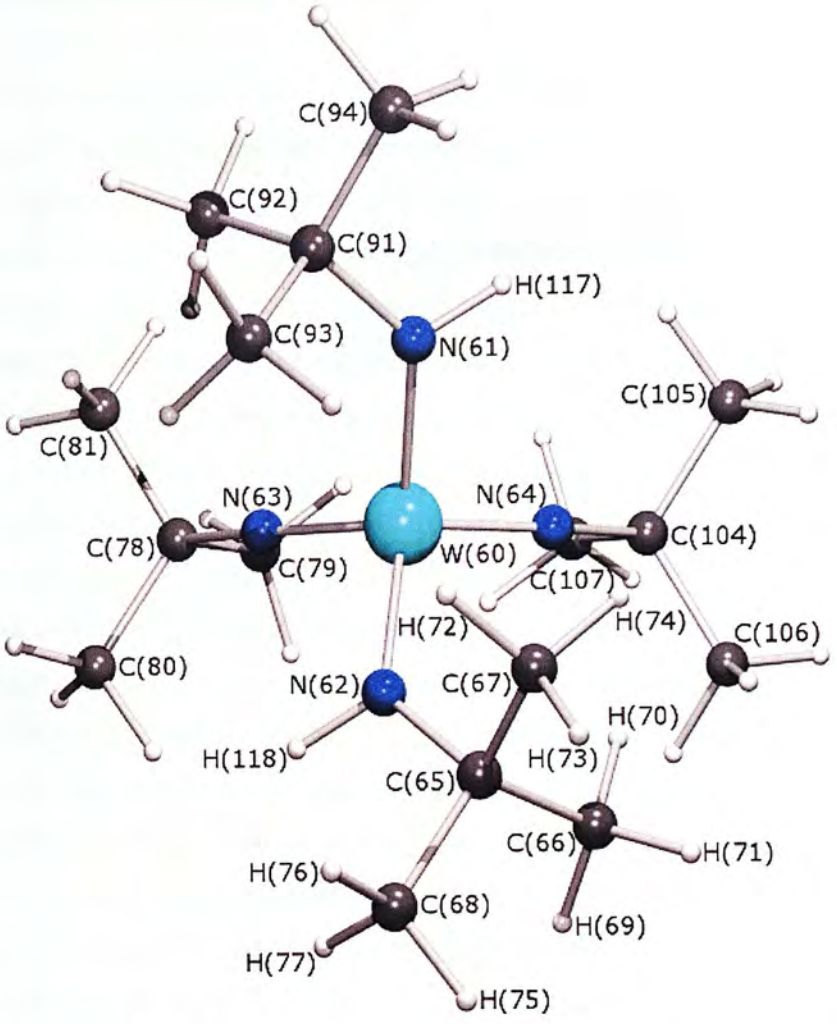
5.2. Experimental

5.2.1. Computational methods

A search of the potential-energy surface located two conformers of $W(NBu^t)_2(NHBu^t)_2$, shown in Figure 5.1, conformer 1 with C_1 symmetry and conformer 2 with C_2 symmetry. The difference in energy between the conformers was calculated to be sufficiently low that a substantial amount of each conformer could be expected in the experimental gas-phase sample mixture. Geometry optimisations were then carried out for both of these conformers as detailed in Chapter 2, using a LanL2DZ pseudopotential[10-12] on the W atom, as Pople-style basis sets beyond 3-21G* do not include third-row transition metals. The amplitudes of vibration (u_{h1}) and curvilinear vibrational correction factors (k_{h1}) to distances were also derived as previously described.



Conformer 1 (C_1 symmetry)



Conformer 2 (C_2 symmetry)

Figure 5.1. Gas-phase molecular structures of conformers 1 and 2 of $W(NBu^t)_2(NHBu^t)_2$. In each case, only one butyl group has all the hydrogen atoms labelled, for simplicity. The full connectivity can be found later in Tables 4.1 and 4.2.

5.2.2. Electron diffraction notes

The Edinburgh gas-phase electron diffraction (GED) apparatus was used to collect data from a sample of $W(NBu^t)_2(NHBu^t)_2$ synthesised in Bath.[13] The scattering factors of Ross *et al*[14] were used in the refinements.

5.2.3. Electron diffraction model

To describe the structure of $W(NBu^t)_2(NHBu^t)_2$ a two-conformer model was written using 42 geometric parameters, comprising seven bond lengths, 20 bond angles and 15 torsions. A non-geometric parameter to control the relative amount of each conformer in the mixture (p_{43}) was also included.

The equivalent bond lengths calculated for the two conformers were almost identical at the B3PW91 level of theory, and it was thus determined that average bond-length parameters could be used for both conformers with fixed differences between them. This led to the set of seven bond-length parameters. The W–N distances were described by the average and a difference (p_{1-2}), defining the difference between the average W–N and the average W–N(H) distances. In the case of the C_1 conformer, there was also a small difference (less than 1 pm) between the individual W–N distances and that was included in the model using a non-refining parameter fixed at 0.9 pm. An average and two difference parameters were used to describe the C–C and N–C distances (p_{3-5}). These were defined as follows: p_3 = {[C–C] + [N(2)–C] + [N(4)–C]} / 3, p_4 = [N(2)–C] – {[N(4)–C] + (C–C)} / 2 and p_5 = [N(4)–C] – (C–C). *Ab initio* calculations showed that while in principle there are three different C–C distances in each *tert*-butyl group, the differences are so small that there would be no benefit in using more than one parameter to describe all 24 of these. Two further distance parameters were used to describe the H-atom positions, *r*C–H mean (p_6) for the methyl hydrogens, and *r*N–H for the two imide hydrogens (p_7).

It was necessary to use different angle parameters for the two separate conformers. For the C_2 conformer (conformer 2), these comprised the average of N(4)–W(1)–N(5) and N(3)–W(1)–N(2) and the difference between them (p_{8-9}), mean values for ∠W–N–C (p_{10}), ∠W–N(H)–C (p_{11}), ∠W–N–H (p_{12}), ∠N–C–C (p_{13}) and finally ∠C–C–H (p_{14}).

For conformer 1 more parameters were required, since the structure exhibits only C_1 symmetry. To describe the N–W–N angles, five parameters were required. The average of the three N(4)–W–N angles ([N(4)–W(1)–N(2) + N(4)–W(1)–N(3) + N(4)–W(1)–N(5)] / 3) was used (p_{21}), along with two differences – the difference between N(4)–W–N(5) and the value of ([N(4)–W–N(2) + N(4)–W–N(3)] / 2) (p_{22}),

and the difference between N(4)–W–N(2) and N(4)–W–N(3) (p_{23}). Two further parameters, an average and a difference, were used for the N(5)–W–N angles, the average of N(5)–W–N(2) and N(5)–W–N(3) (p_{24}), and the difference between them (p_{25}). The average of the two W–N(H)–C angles and the difference between them were used (p_{28} and p_{29}), and also the average and difference of the two W–N–C angles (p_{26} and p_{27}). The remaining angle parameters used for the C_1 conformer were the average and difference ∠W–N–H (p_{30} and p_{31}), and mean values of ∠N–C–C (p_{32}) and ∠C–C–H (p_{33}).

Different torsional parameters were also required for the two different conformers. In the case of the C_2 conformer, these were a single methyl-group torsion (p_{18}), parameters to describe the torsions of the butyl groups within the NBu^t ligands (p_{16}) and within the NHBu^t ligands (p_{17}), and a torsion of the overall NHBu^t groups about the N–W bond (ϕH–N–W–N; p_{15}). In order to model the N atoms about the central tungsten atom, a further parameter was used, defined as a torsion about the y axis (p_{20}), where X is defined as the point (0.000, 1.000, 0.000), that is, a point exactly 100 pm from the W atom (which sits at the origin), along the y axis. A parameter to describe the drop of the butyl groups out of the W–N–H plane was also included (p_{19}). All torsion angles were defined to have the same signs for rotations of the respective groups in the same sense.

For the C_1 conformer, again, more parameters were required. The four ligands to tungsten were labelled groups 1–4, with group 1 including N(4), group 2 including N(5), group 3 including N(3) and group 4 including N(2). A single methyl-group torsion was used (p_{40}), and, in this case, individual torsions for each of the four butyl groups (p_{36}, p_{37}, p_{38} and p_{39}). Individual torsions of the two NHBu^t groups (groups 3 and 4) about the N–W bonds were also used (p_{34} and p_{35}), along with parameters to describe the out-of-plane drop of each of these groups (p_{41} and p_{42}).

Using the Boltzmann equation, and taking into account that conformer 1 has double multiplicity, it was calculated that the relative amounts should be 70% of conformer 2, with 30% conformer 1. The proportionality parameter p_{43} was therefore fixed at 70% for the initial least-squares refinement.

5.3. Results

5.3.1. *Ab initio* calculations

The presence of a heavy transition-metal atom (W) in a molecule normally results in pure *ab initio* methods performing poorly, and that is observed here. Table 5.1 illustrates that the geometries show considerable variation between the MP2 and B3PW91 levels of theory. The geometry of conformer 2 calculated using B3PW91//LanL2DZ/6-31G* has a W–N bond length that is shorter by 2.7 pm than that calculated using MP2, with the W–N(H) bond shorter by 1.1 pm. A related effect can also be seen in the calculated geometries of conformer 1, where the W–N distances are also shortened by 2.7 pm, although this time there is no significant change to the W–N(H) distances. The N–W–N angles are also affected by the method used. For example, ∠N(63)–W(60)–N(64) increases from 113.8° to 115.9° moving from MP2 to B3PW91, and the overall range of N–W(60)–N angles increases by 2.5°. For both conformers, the only parameters significantly affected by alteration of the level of theory were those involving the tungsten atom.

Table 5.1. Comparison of heavy-atom parameters for $W(NBu^t_3)_2(NHBu^t_3)_2$ calculated using the MP2 and B3PW91 methods. The 6-31G* basis set was used on all atoms except W, for which the LanL2DZ pseudopotential was used. All distances are in pm and all angles are in degrees.

Parameter	MP2	B3PW91
*r*W(60)–N	176.6	179.5
*r*W(60)–N(H)	198.6	197.4
*r*W(1)–N(4)	174.2	176.9
*r*W(1)–N(5)	173.6	176.1
*r*W(1)–N(2)	197.6	197.5
*r*W(1)–N(3)	197.7	197.6

∠N(63)–W(60)–N(64)	113.8	115.9
∠N(63)–W(60)–N(61)	110.3	111.2
∠N(63)–W(60)–N(62)	104.2	103.8
∠N(4)–W(1)–N(5)	115.2	115.9
∠N(4)–W(1)–N(2)	109.9	110.6
∠N(4)–W(1)–N(3)	109.8	109.6
∠N(5)–W(1)–N(2)	103.5	103.3
∠N(5)–W(1)–N(3)	105.9	105.4
∠N(2)–W(1)–N(3)	111.7	112.6

The Gibbs free energies of the conformers were calculated at the B3PW91/LanL2DZ level of theory with two basis sets, 6-31G* and 6-311+G*. At the 6-31G* level, conformer 2 was found to be lower in energy by ~10 kJ mol^{-1}, while increasing the basis set to 6-311+G* altered this greatly, giving conformer 2 lower in energy by ~5.5 kJ mol^{-1}. It was, therefore, possible that both conformers would be present in the experimental mixture, and so both were included in the model written for the least-squares refinement of the GED data. However, the instability of the calculated energy difference with respect to basis set meant that the absolute value could not be trusted to give an accurate prediction of the conformational make-up.

The values of the peripheral-atom parameters calculated at the MM and MP2 levels of theory are given in Table 5.2 for one *tert*-butyl group of each conformer. The full set of peripheral-atom parameters is given in Table S5.1 in the Supplementary Information.

Table 5.2. Comparison of calculated parameters for $W(NBu^t_3)_2(NHBu^t_3)_2$ calculated using the MM3 and B3PW91 (DFT) methods. The 6-311+G* basis set was used on all atoms except W, for which the LanL2DZ pseudopotential was used. Only one *tert*-butyl group is shown for each conformer. The full set of peripheral-atom parameters can be found in the Supplementary Information. All distances are in pm and all angles are in degrees.

Parameter	*r*C-H		∠C-C-H		ϕN-C-C-H	
Conformer 1	MM3	DFT	MM3	DFT	MM3	DFT
H(12)-C(9)-C(8)-N(2)	111.1	109.4	112.7	110.3	-66.9	-60.9
H(13)-C(9)-C(8)-N(2)	111.3	109.3	112.0	110.3	55.0	59.1
H(14)-C(9)-C(8)-N(2)	111.2	109.4	111.1	110.7	174.2	179.5
H(15)-C(10)-C(8)-N(2)	111.2	109.6	112.0	110.9	60.7	59.6
H(16)-C(10)-C(8)-N(2)	111.2	109.4	111.6	110.6	-179.5	179.1
H(17)-C(10)-C(8)-N(2)	111.3	109.4	111.6	110.7	-59.6	-60.6
H(18)-C(11)-C(8)-N(2)	111.2	109.6	111.4	111.9	-179.3	178.4
H(19)-C(11)-C(8)-N(2)	111.3	109.4	111.7	110.3	-59.7	-61.1
H(20)-C(11)-C(8)-N(2)	110.9	109.1	112.4	109.5	60.7	57.8
Conformer 2	MM3	DFT	MM3	DFT	MM3	DFT
H(69)-C(66)-C(65)-N(62)	111.3	109.5	112.3	110.9	-50.6	-60.2
H(70)-C(66)-C(65)-N(62)	110.8	109.3	112.0	110.1	71.8	59.1
H(71)-C(66)-C(65)-N(62)	111.2	109.6	110.6	111.6	-170.0	179.6
H(72)-C(67)-C(65)-N(62)	111.2	109.3	111.8	110.7	57.8	58.0
H(73)-C(67)-C(65)-N(62)	111.2	109.4	110.7	110.7	176.6	178.2
H(74)-C(67)-C(65)-N(62)	111.1	109.3	113.1	111.0	-64.0	-61.9
H(75)-C(68)-C(65)-N(62)	111.2	109.5	110.8	111.2	-178.4	179.4

Parameter	rC-H		∠C-C-H		ϕN-C-C-H	
H(76)-C(68)-C(65)-N(62)	111.3	109.3	112.2	110.8	-59.2	-60.2
H(77)-C(68)-C(65)-N(62)	111.3	109.6	112.4	111.1	62.3	59.9

5.3.2. SARACEN refinement

The starting values for the 42 geometric parameters used in the refinement were taken from the highest-level DFT calculation carried out (B3PW91//LanL2DZ/6-311+G*). The model was refined as an r_{h1} structure (using curvilinear amplitude corrections). In total, all 42 geometric parameters and 10 groups of amplitudes were refined. Where parameters were not sufficiently well defined to refine to sensible values restraints were applied in accordance with the SARACEN method.

After all parameters and amplitude groups were refined, the parameter describing the amount of C_2 conformer was allowed to change to find the value that best fitted the experimental data. This was done by incrementally stepping through different conformer amounts to find the minimum R_G value. The final R factor for the refinement was $R_G = 0.068$ ($R_D = 0.046$), which was obtained for a conformer ratio of 69.5:30.5 (C_1:C_2). Figure 5.2 shows the R-factor ratio, R_G/R_G(min.), plotted against the percentage of conformer 2. As can be seen, the R factor increases quite sharply either side of the minimum at 30.5%. The 95% confidence limit[15] [where R_G/R_G(min.) = 1.016] shows that the uncertainty in this value is less than ±1%.

Final refined parameters are listed in Table 5.3. Interatomic distances and the corresponding amplitudes of vibration are given in Table S5.2, with the final experimental coordinates for the two-conformer GED analysis given in Table S5.3. The least-squares correlation matrix is given in Table S5.4.

5.3.3. DYNAMITE refinement

The starting values for the 42 geometric parameters used in the refinement were taken from the completed SARACEN refinement. As before, all 42 geometric

parameters and 10 groups of amplitudes were refined, with the same restraints on parameters applied as for the SARACEN method. The final R factor for the refinement was $R_G = 0.068$ ($R_D = 0.046$), which was obtained for the same conformer ratio of 69.5:30.5 (C_1:C_2) as for the SARACEN method.

Final refined parameters are listed in Table 5.3, with peripheral-atom parameters for one *tert*-butyl group of each conformer in Table 5.4. Interatomic distances and the corresponding amplitudes of vibration are given in Table S5.5, with the final experimental coordinates for the two-conformer GED analysis given in Table S5.6. The least-squares correlation matrix is given in Table S5.7.

5.3.4. SEMTEX refinement

As for the DYNAMITE study, the starting values for the 42 geometric parameters used in the refinement were taken from the completed SARACEN refinement. As before, all 42 geometric parameters were refined along with 10 groups of amplitudes, with the same restraints on parameters applied as for the SARACEN method. The final R factor for the refinement was $R_G = 0.068$ ($R_D = 0.046$), which was again obtained for the same conformer ratio of 69.5:30.5 (C_1:C_2).

Final refined parameters are listed in Table 5.3, with peripheral-atom parameters for one *tert*-butyl group of each conformer given in Table 5.4. The radial-distribution curve for the refinement is shown in Figure 5.3, and Figure 5.4 shows the molecular-scattering intensity curve from the refinement. Interatomic distances and the corresponding amplitudes of vibration are given in Table S5.8, with the final experimental coordinates for the two-conformer GED analysis given in Table S5.9. The least-squares correlation matrix is given in Table S5.10. The full set of peripheral-atom parameters can be found in Table S5.11 in the Supplementary Information.

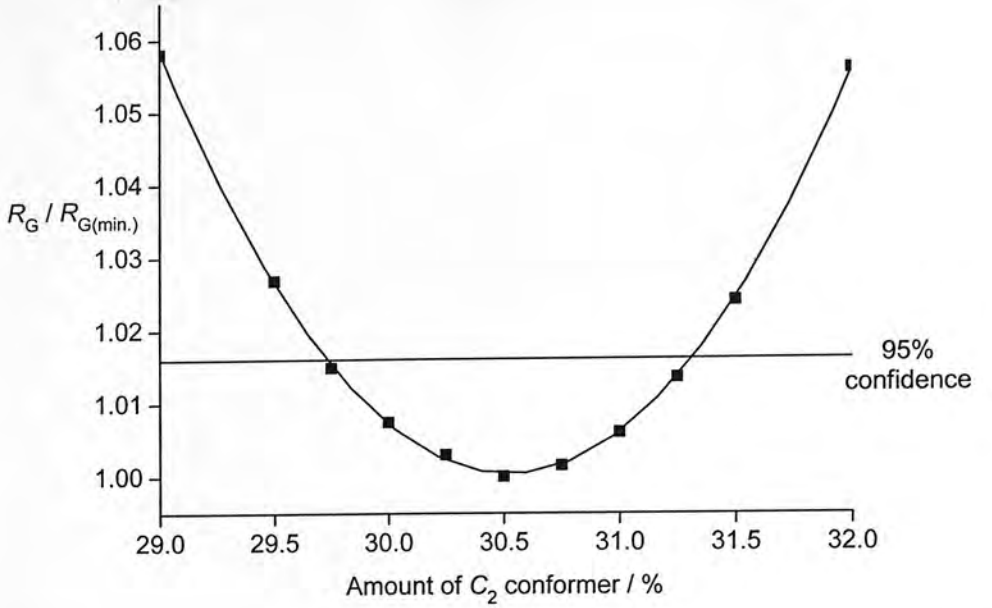


Figure 4.2. R_G / R_G min. for varying amounts of conformer 2 in the SEMTEX refinement of $W(NBu^t)_2(NHBu^t)_2$.

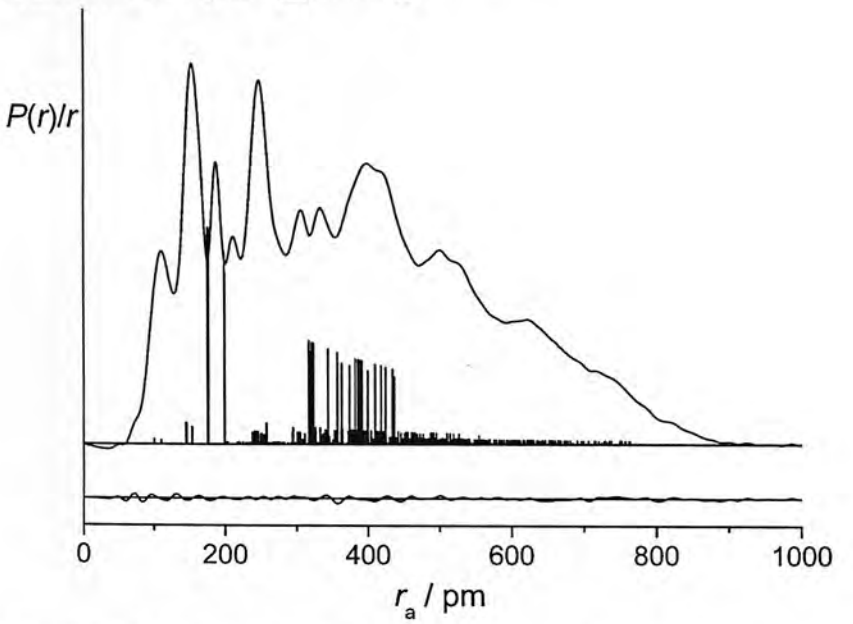


Figure 5.3 – Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $W(NBu^t)_2(NHBu^t)_2$. Before Fourier inversion the data were multiplied by $s\cdot\exp(-0.00002s^2)/(Z_N-f_N)(Z_W-f_W)$.

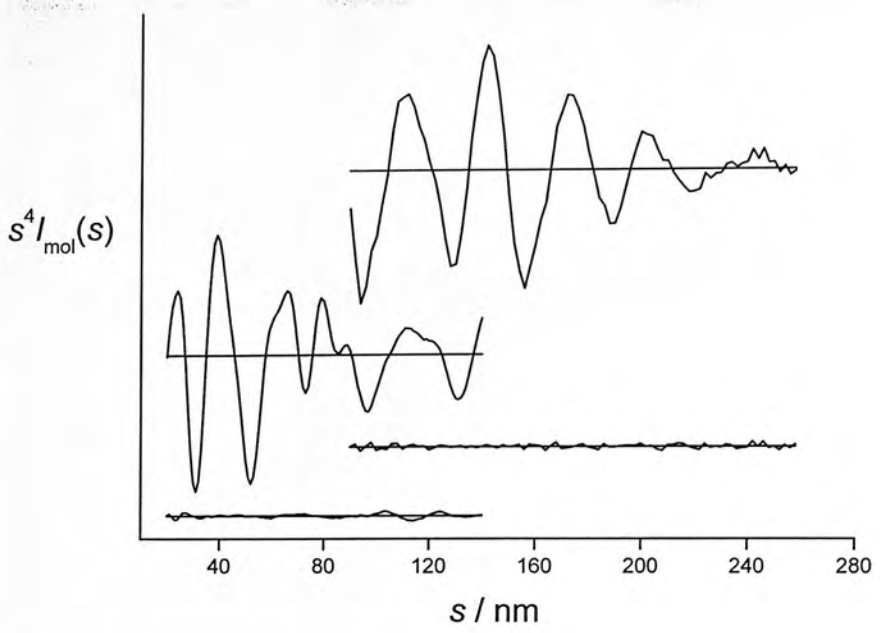


Figure 5.4. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $W(NBu^t)_2(NHBu^t)_2$.

Table 5.3. Refined and calculated parameters for the two-conformer SARACEN refinement of $W(NBu^t)_2(NHBu^t)_2$. Distances are in pm and angles in degrees. The LanL2DZ pseudopotential was used for W and the 6-31G* basis set for H, C and N. X is a dummy atom lying in the positive direction on the x-axis.

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
Independent						
Both conformers						
p_1	rW–N average	187.2	188.0(1)	188.0(1)	188.0(1)	
p_2	rW–N difference	20.6	23.3(4)	23.3(5)	23.3(4)	
p_3	rC–C/N–C average	148.1	148.4(2)	148.2(2)	148.4(2)	
p_4	rC–C/N–C difference 1	–1.9	–3.9(8)	–3.7(8)	-3.9(8)	–1.9(10)
p_5	rC–C/N–C difference 2	–9.4	–8.2(5)	–8.7(6)	-8.3(5)	–9.4(10)
p_6	rC–H mean	109.0	110.9(2)	110.9(2)	110.8(2)	
p_7	rN–H average	101.3	101.3(5)	101.3(5)	101.3(5)	101.3(5)

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
Conformer 2						
p_8	∠N–W–N average conf2	113.5	113.5(9)	113.4(9)	113.4(10)	
p_9	∠N–W–N difference conf2	4.9	5.0(10)	5.0(10)	5.0(10)	4.9(10)
p_{10}	∠W–N–C average conf2	163.6	164.0(10)	164.1(10)	164.0(10)	
p_{11}	∠W–N(H)–C average conf2	133.4	133.4(10)	133.6(10)	133.4(10)	133.4(10)
p_{12}	∠W–N–H conf2	114.7	114.8(10)	114.8(10)	114.8(10)	114.7(10)
p_{13}	∠N–C–C average conf2	109.2	106.7(7)	106.6(7)	106.8(7)	
p_{14}	∠C–C–H average conf2	109.5	110.0(12)	110.0(9)	110.1(12)	109.5(15)
p_{15}	ϕH–N–W–N conf2	114.6	117.0(41)	117.1(40)	117.1(42)	114.6(50)
p_{16}	Bu^t twist (NBu^t groups) conf2	18.0	61.9(49)	62.0(51)	62.2(51)	
p_{17}	Bu^t twist ($NHBu^t$ groups) conf2	29.7	61.1(77)	60.1(80)	60.5(80)	

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{18}	Me twist conf2	180.0	160.0(42)	159.1(43)	160.0(42)	
p_{19}	Bu^t drop conf2	13.5	13.4(10)	13.4(10)	13.4(10)	
p_{20}	ϕN–W–X–N[c] conf2	95.1	96.1(18)	96.0(20)	96.2(19)	
Dependent parameters						
d_1	rW(60)-N	179.5	176.3	176.3	176.4	
d_2	rW(60)-N(H)	197.4	199.6	199.6	199.6	
d_3	∠N(63)–W(60)–N(64)	115.9	115.7	115.9	115.9	
d_4	∠N(63)–W(60)–N(61)	111.2	112.2	112.0	112.1	
d_5	∠N(63)–W(60)–N(62)	103.8	103.0	103.0	103.0	
Conformer 1						
p_{21}	∠N(4)–W–N average conf1	111.8	111.4(6)	111.5(6)	111.4(6)	
p_{22}	∠N(4)–W–N difference conf1	6.1	7.6(9)	7.8(9)	7.7(9)	6.1(10)

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{23}	∠N(4)–W–N(H) difference conf1	0.4	0.4(1)	0.4(1)	0.4(1)	0.4(1)
p_{24}	∠N(5)–W–N average conf1	104.3	104.5(8)	104.5(8)	104.6(8)	
p_{25}	∠N(5)–W–N difference conf1	2.1	2.2(10)	2.3(10)	2.3(10)	2.1(10)
p_{26}	∠W–N–C average conf1	163.7	163.5(7)	163.5(7)	163.5(7)	163.5(10)
p_{27}	∠W–N–C difference conf1	2.5	2.3(10)	2.3(10)	2.3(10)	2.5(10)
p_{28}	∠W–N(H)–C average conf1	136.6	135.1(8)	135.2(8)	135.0(8)	136.6(10)
p_{29}	∠W–N(H)–C difference conf1	3.0	2.9(10)	2.9(10)	3.0(10)	3.0(10)
p_{30}	∠W–N–H average conf1	112.5	112.6(10)	112.6(10)	112.6(10)	112.5(10)
p_{31}	∠W–N–H difference conf1	2.0	2.0(5)	2.0(5)	2.0(5)	2.0(5)
p_{32}	∠N–C–C mean conf1	109.2	109.3(4)	109.5(4)	109.2(4)	109.2(10)
p_{33}	∠C–C–H mean conf1	109.5	112.0(4)	111.9(9)	112.0(4)	

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{34}	Twist NHBu^t group3 conf1	149.9	151.1(48)	151.3(46)	151.1(47)	149.9(50)
p_{35}	Twist NHBu^t group4 conf1	–135.8	–133.4(36)	-134.0(36)	-133.8(36)	–135.8(50)
p_{36}	Bu^t twist group1 conf1	23.0	30.0(29)	30.5(29)	30.1(29)	23.0(50)
p_{37}	Bu^t twist group2 conf1	18.0	19.3(30)	19.3(29)	19.1(30)	18.0(30)
p_{38}	But twist group3 conf1	–147.5	–140.8(53)	-140.0(50)	-140.4(52)	-147.5(100)
p_{39}	Bu^t twist group4 conf1	32.3	40.9(45)	40.6(44)	41.5(46)	32.3(100)
p_{40}	Me twist conf1	180.0	179.5(46)	179.7(45)	179.4(46)	
p_{41}	Bu^t drop group3 conf1	15.0	15.2(21)	15.3(22)	15.3(21)	15.0(20)
p_{42}	Bu^t drop group4 conf1	12.8	12.5(20)	12.4(20)	12.6(20)	12.8(20)
p_{43}	Percentage of conf2[a]	56	30.5(10)	30.5(10)	30.5(10)	
Dependent parameters						
d_6	rW(1)–N(4)	176.9	176.9	176.8	176.9	

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	DYNAMITE (r_{h1})	SEMTEX (r_{h1})	Restraint
d_7	rW(1)–N(5)	176.1	175.8	175.8	175.9	
d_8	rW(1)–N(2)	197.5	199.7	199.7	199.6	
d_9	rW(1)–N(3)	197.6	197.6	199.7	199.6	
d_{10}	∠N(4)–W(1)–N(5)	115.9	116.7	116.7	116.6	
d_{11}	∠N(4)–W(1)–N(2)	110.6	109.2	109.1	109.1	
d_{12}	∠N(4)–W(1)–N(3)	109.6	108.7	108.7	108.7	
d_{13}	∠N(5)–W(1)–N(2)	103.3	103.3	103.3	103.4	
d_{14}	∠N(5)–W(1)–N(3)	105.4	105.7	105.6	105.7	
d_{15}	∠N(2)–W(1)–N(3)	112.6	113.5	113.4	113.5	

[a] This parameter was not refined, but rather fixed at the value calculated using the Boltzmann equation at the experimental temperature (56%) whilst the refinement of geometric parameters was in progress. Once this was complete, p_{43} was systematically set to different values and the resulting R factors plotted in order to find the optimum conformer mixture.

Table 5.4. Comparison of refined peripheral-atom parameters for one *tert*-butyl group of each conformer of $W(NBu^t_3)_2(NHBu^t_3)_2$, obtained using the DYNAMITE and SEMTEX methods. The full set of peripheral-atom parameters can be found in the Supplementary Information. All distances are in pm and all angles are in degrees.

Parameter	*r*C-H		∠C-C-H		ϕN-C-C-H	
Conformer 1	DYN	SEM	DYN	SEM	DYN	SEM
H(12)-C(9)-C(8)-N(2)	110.8	110.8	111.0	111.8	-60.7	-60.5
H(13)-C(9)-C(8)-N(2)	111.0	110.8	110.3	111.7	61.2	61.6
H(14)-C(9)-C(8)-N(2)	110.9	110.8	109.4	112.1	-179.6	-179.3
H(15)-C(10)-C(8)-N(2)	110.9	111.0	110.3	112.3	60.5	60.8
H(16)-C(10)-C(8)-N(2)	111.0	110.8	109.9	112.0	-179.7	-179.4
H(17)-C(10)-C(8)-N(2)	111.0	110.9	109.9	112.1	-59.8	-59.6
H(18)-C(11)-C(8)-N(2)	110.9	111.1	109.7	113.3	-179.6	-179.3
H(19)-C(11)-C(8)-N(2)	110.9	110.9	110.0	111.7	-59.9	-59.7
H(20)-C(11)-C(8)-N(2)	110.6	110.6	110.7	110.9	60.4	60.8
Conformer 2	DYN	SEM	DYN	SEM	DYN	SEM
H(69)-C(66)-C(65)-N(62)	110.9	111.0	111.9	110.2	-81.9	-81.0
H(70)-C(66)-C(65)-N(62)	110.5	110.8	111.5	109.5	40.5	41.4
H(71)-C(66)-C(65)-N(62)	110.9	111.0	111.9	111.0	158.6	159.7
H(72)-C(67)-C(65)-N(62)	110.9	110.8	111.9	109.6	40.1	41.0
H(73)-C(67)-C(65)-N(62)	110.9	110.9	111.9	109.6	159.0	159.9
H(74)-C(67)-C(65)-N(62)	110.8	110.8	111.8	110.0	-81.7	-80.8
H(75)-C(68)-C(65)-N(62)	110.9	110.9	111.9	110.5	159.1	160.0

Parameter	rC-H		∠C-C-H		ϕN-C-C-H	
H(76)-C(68)-C(65)-N(62)	111.0	110.8	112.0	110.0	-81.6	-80.7
H(77)-C(68)-C(65)-N(62)	111.0	111.1	112.0	110.4	40.1	40.8

5.4. Discussion

The molecular structure of $W(NBu^t)_2(NHBu^t)_2$ in the gas phase has been determined by GED. The refinement was first carried out using the SARACEN method, and then the result used as a starting geometry for DYNAMITE and SEMTEX refinements.

For this molecule, the *tert*-butyl groups are rather far apart both from one another and from the central tungsten atom, and steric effects are therefore limited. This means that there is little distortion away from C_3 symmetry of the *tert*-butyl groups, and consequently that neither of the DYNAMITE or SEMTEX methods made any significant difference to the final refined structure for this molecule. Looking at Table 5.3, it can be seen that with both the DYNAMITE and SEMTEX methods, the range of light-atom distances and angles is small, with only a ~2.0° variation in angles for each conformer. The near-identical R-factor obtained upon using SEMTEX or DYNAMITE instead of SARACEN, and the closeness of the methyl groups to C_3 symmetry, show that these methods are not artificially influencing the refinement and that the SEMTEX code has been successfully developed to deal with multiple-conformer refinements.

The two-conformer model gave a much better fit to the electron-diffraction data than either of two single-conformer models, vindicating the computed result that both conformers were present in significant quantities in the experimental sample. The presence of roughly twice as much C_1 conformer as C_2 is a consequence of the fact that the C_1 conformer is doubly-degenerate (it exists in two forms which are identical to GED analysis), whilst the C_2 conformer exhibits no degeneracy, and so indicates that the two conformers are of approximately equal energy. The calculated energy difference of approximately +5.5 kJ mol^{-1} (with the positive sign defined as showing that the C_2 symmetry conformer is lower in energy) indicated a mixture comprising

70% of conformer 2 and 30% conformer 1. This is therefore very different from the final conformational composition obtained from the refinement, which indicates a small energy difference of –0.47(7) kJ mol^{-1} (C_1 symmetry conformer lower in energy) between the conformers, and a mixture of 69.5% conformer 1 and 30.5% conformer 2. The uncertainty in the energy difference between conformers was estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 2 (see Figure 5.2). However, the calculated energy difference is very unstable with respect to the basis set, and so can only reliably be used to indicate whether both conformers are likely to be present.

For both conformers, the refined distances and angles are in reasonably good agreement with those calculated using the B3PW91//LanL2DZ/6-311+G* method. The largest discrepancy occurs with the C–C/N–C difference parameter (p_4), which increases from –1.9 pm to –3.9 pm. This has the effect of bringing the values of the N–C and N(H)–C distances (dependent parameters) much closer together [145.5(8) pm and 145.8(5) pm, compared to 144.0 pm and 146.8 pm]. The largest difference in angle occurs with the NCC angle in the C_1 symmetry conformer, which refines to 106.7°, compared with the computed value of 109.2°.

Torsional parameters are also generally in good agreement with computed values. However, the torsion angles of the *tert*-butyl groups in the C_2-symmetric conformer differ significantly: 61.9(49)° compared to the calculated 18.0° for the NBu^t groups, and 61.1(77)° compared to 29.7° for the NHBu^t groups.

The W–N distances, which make the biggest single contribution to the overall scattering intensity, are in generally good agreement with the values calculated at the B3PW91 level of theory. The most notable difference occurs in the longer W–N(H) distance, which refines to 199.6 pm for both conformer 1 and conformer 2, compared with the calculated values of 197.4 pm for conformer 1 and 197.5 pm for conformer 2. The refined value is considerably closer to the sum of the covalent radii of tungsten and nitrogen (200.0 pm). The average W–N double-bonded distance refines to 176.4 pm, compared with calculated values of 176.8 pm for both conformers.

There is a sharp increase in R_G on each side of the optimal value for relative conformer amounts (amount of C_2 conformer = 30.5%), indicating that the

conformational make-up of the sample mixture is well defined. That the two conformers lie so close in energy is not surprising when their relative structures are considered – the only significant difference between the C_1 and C_2 symmetric forms is in the conformation of a single $NHBu^t$ group, which has only a very small contribution to the overall energy of the molecule.

5.5. Conclusion

The molecular structure of $W(NBu^t)_2(NHBu^t)_2$ in the gas phase has been determined by GED and used as a further test case for the new SEMTEX method. The refinement indicated a conformational composition of 30.5% C_2 conformer and 69.5% C_1 conformer, a notably different result to that calculated using Hartree-Fock methods. Neither the DYNAMITE nor SEMTEX methods made any discernable improvement to the structure for this molecule, since *tert*-butyl groups experience very little distortion away from local C_3 symmetry, as indicated by *ab initio* calculations. However, the fact that the SEMTEX refinement provides a near-identical result to the SARACEN and DYNAMITE structures in this case shows that the method is now being correctly applied to multi-conformer systems and is not having an artificial influence on the structure when none is warranted. This indicates that the method is now suitable for further use on more challenging systems where it could be expected to have an effect on the overall quality of refinement.

5.6. References

1. W. A. Nugent and B. L. Haymore, *Coord. Chem. Rev.* 1980, **31**, 123.
2. W. A. Nugent, *Inorg. Chem.* 1983, **22**, 965-9.
3. W. A. Nugent and R. L. Harlow, *Inorg. Chem.* 1980, **19**, 777-9.
4. A. A. Danopoulos, G. Wilkinson, B. Hussain and M. B. Hursthouse, *J. Chem. Soc., Chem. Commun.* 1989, 896-7.
5. A. A. Danopoulos, G. Wilkinson, B. Hussain-Bates and M. B. Hursthouse, *J. Chem. Soc., Dalton Transactions* 1990, 2753.
6. A. A. Danopoulos, G. Wilkinson, T. K. N. Sweet and M. B. Hursthouse, *Polyhedron* 1994, **13**, 2899.
7. E. L. Crane, H.-T. Chiu and R. G. Nuzzo, *J. Phys. Chem. B* 2001, **105**, 3549-3556.
8. M. H. Tsai, S. C. Sun, H.-T. Chiu and S.-H. Chuang, *Appl. Phys. Lett.* 1996, **68**, 1412-14.
9. H.-T. Chiu and S.-H.Chuang, *J. Mater. Res.* 1996, **8**, 1353.
10. P. J. Hay and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 270.
11. W. R. Wadt and P. J. Hay, *J. Chem. Phys.* 1985, **82**, 284.
12. P. J. Hay, and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 299.
13. H. Choujaa, S. D. Cosham, A. L. Johnson, G. R. Kafka, M. F. Mahon, S. L. Masters, K. C. Molloy, D. W. H. Rankin, H. E. Robertson and D. A. Wann. *Inorg. Chem.*, 2009, **48**, 2289-2299.
14. A. W. Ross, M. Fink, R. Hilderbrand, Ed. A. J. C. Wilson, Kluwer Academic Publishers: Dordrecht, Boston and London, 1992; **Vol. C**.
15. W. C. Hamilton, *Acta Crystallogr.* **1965**, *18*, 502.

6. Comparison of the gas-phase molecular structures of $C(SiMe_2Cl)_4$, $C(SiMe_2Br)_4$ and $C(SiMe_2F)_4$ determined using each of the SARACEN and SEMTEX methods

6.1. Introduction

With the SEMTEX method now successfully applied to both hydrogen and heavier peripheral atoms, and also used in refinements where multiple conformers were present, it was decided that it could reliably be used in more challenging cases. The next step, therefore, was to find a molecule whose structure would exhibit several of those features that present significant challenges to standard GED, and compare the results of refinements of such a structure using both the SARACEN and SEMTEX methods.

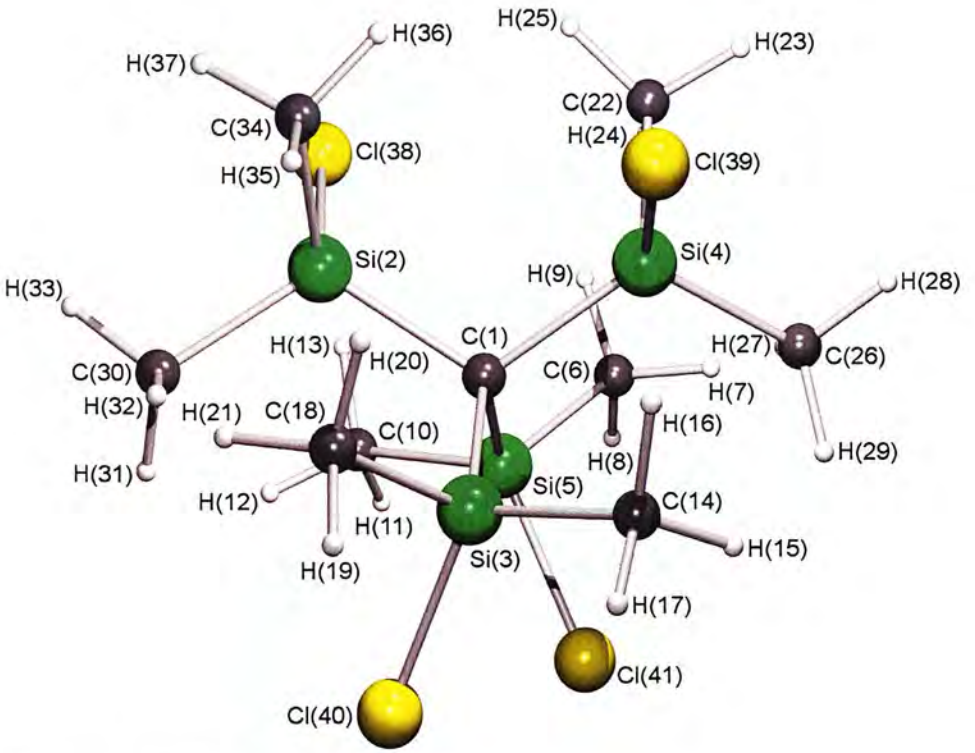
The family of tetrasilylmethane derivatives $C(SiXMe_2)_4$ (X = F, Cl, Br)[1] presents a case where two different types of peripheral atoms are present in the same structure (X and H). This therefore already represents a feature not tackled thus far using SEMTEX. The structure of the related permethyl molecule $C(SiMe_3)_4$ has been well-characterised in the gas-phase using both GED,[2] computational[3,4] and vibrational spectroscopic[4] methods, but little is known about structures of these halogenated variants. Furthermore, an initial search of the potential-energy surface for the chlorinated molecule, $C(SiClMe_2)_4$, revealed the probable presence of several conformers in the GED sample, increasing the likelihood that SEMTEX could provide a tangible improvement in the quality of refinement over SARACEN. Previous NMR work[5] had also indicated several conformers in solution, with possible interconversion between them at higher temperatures, while gas phase electron diffraction studies have also been carried out on the related molecules $(Me_3Si)_3CSiCl_3$[6] and $(HMe_2Si)_3CSiH_3$,[7] in each case showing the presence of eleven distinct conformers.

This work is concerned with investigation of the structure of tetrakis-chlorodimethylsilylmethane using the SARACEN and SEMTEX methods. The SEMTEX code has been altered to enable it to handle two separate sets of peripheral atoms (in this case, X and H) within the same structure. In this work, the SARACEN refinement was initially completed, and results then used as starting geometries for the structural analysis using the SEMTEX method.

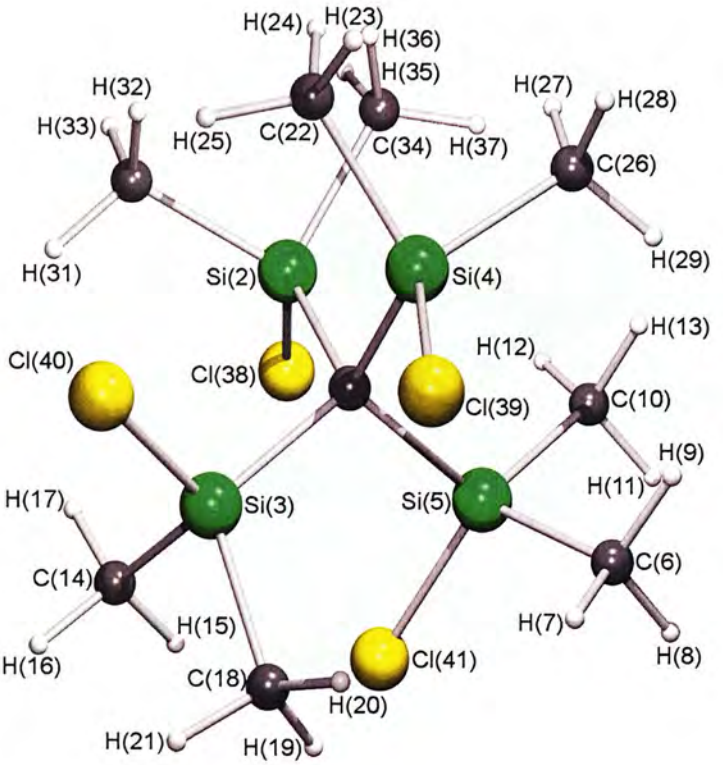
6.2. Experimental

6.2.1. Computational methods

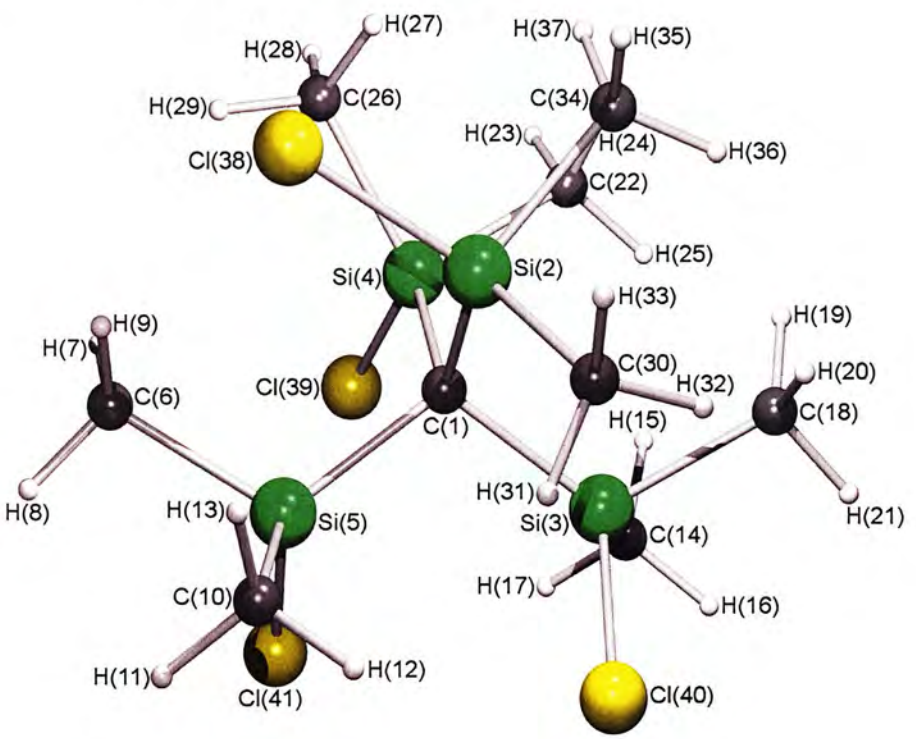
A search of the potential-energy surface located seven potential conformers for each molecule, which were then optimised in each case at the HF/6-31G* level to obtain relative energies. In the case of $C(SiBrMe_2)_4$, a LanL2DZ pseudopotential[8-10] was used on the bromine atoms. Initial calculations using the Boltzmann distribution showed that, for all three molecules, the differences in energies were such that only three of these conformers could be present in the experimental mixture in significant quantities. These were conformer 1, with C_2 symmetry, and conformers 2 and 3, with C_1 symmetry. Geometry optimisations were then carried out for these three conformers. The lowest-energy structures are shown in Figures 6.1, 6.2 and 6.3.



Conformer 1 (C_2 symmetry)

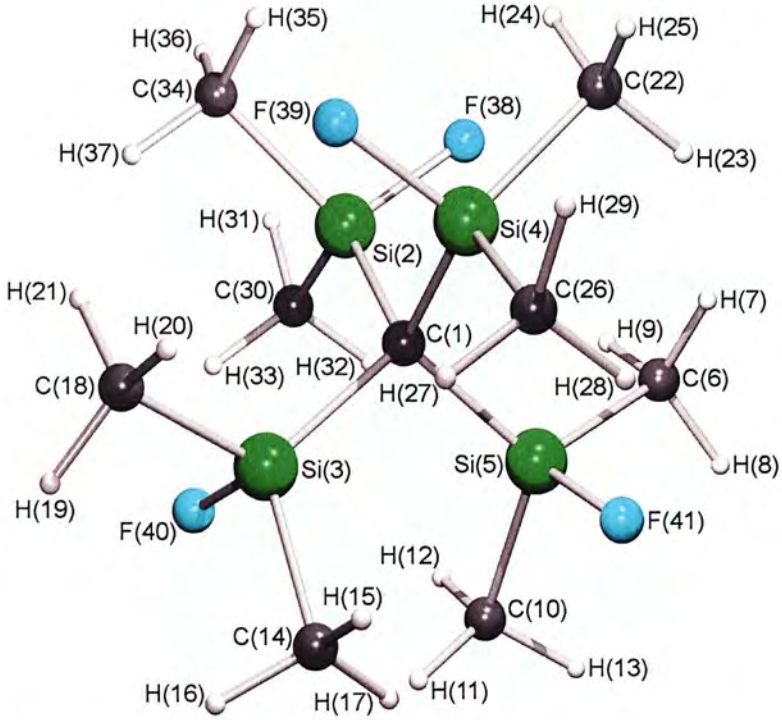


Conformer 2 (C_1 symmetry)

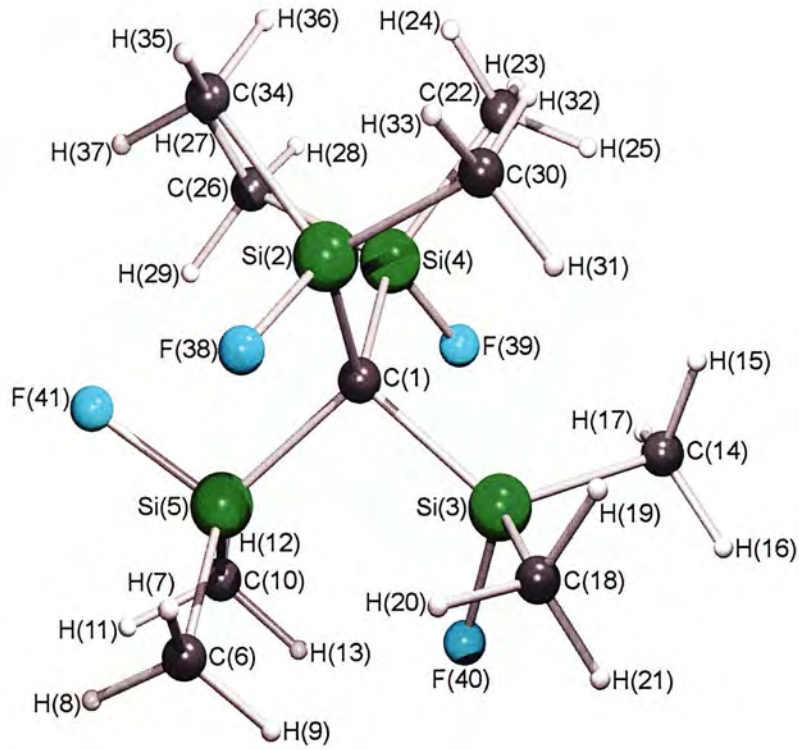


Conformer 3 (C_1 symmetry)

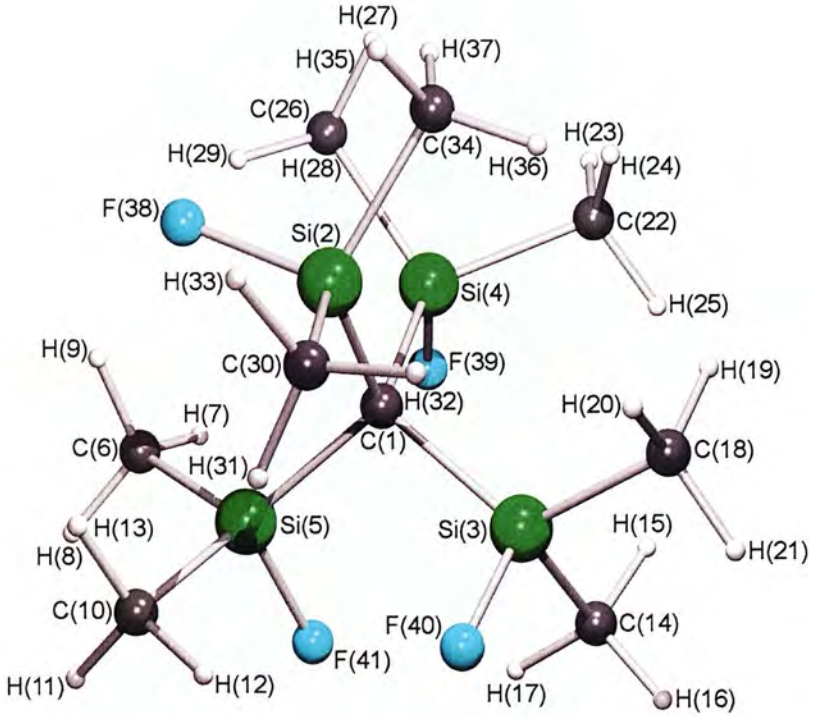
Figure 6.1. Gas-phase molecular structures of conformers 1, 2 and 3 of $C(SiClMe_2)_4$.



Conformer 1 (C_2 symmetry)

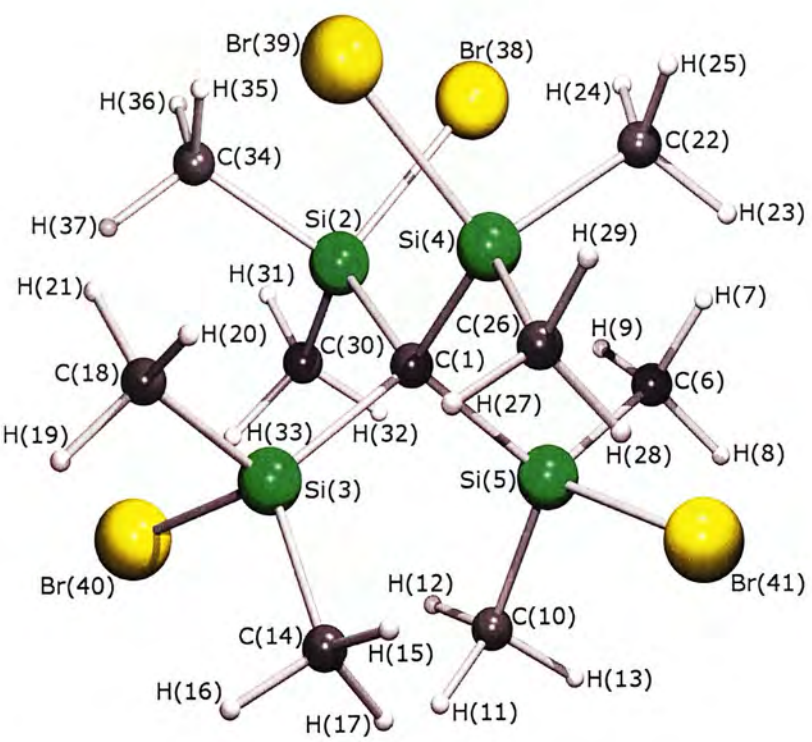


Conformer 2 (C_1 symmetry)

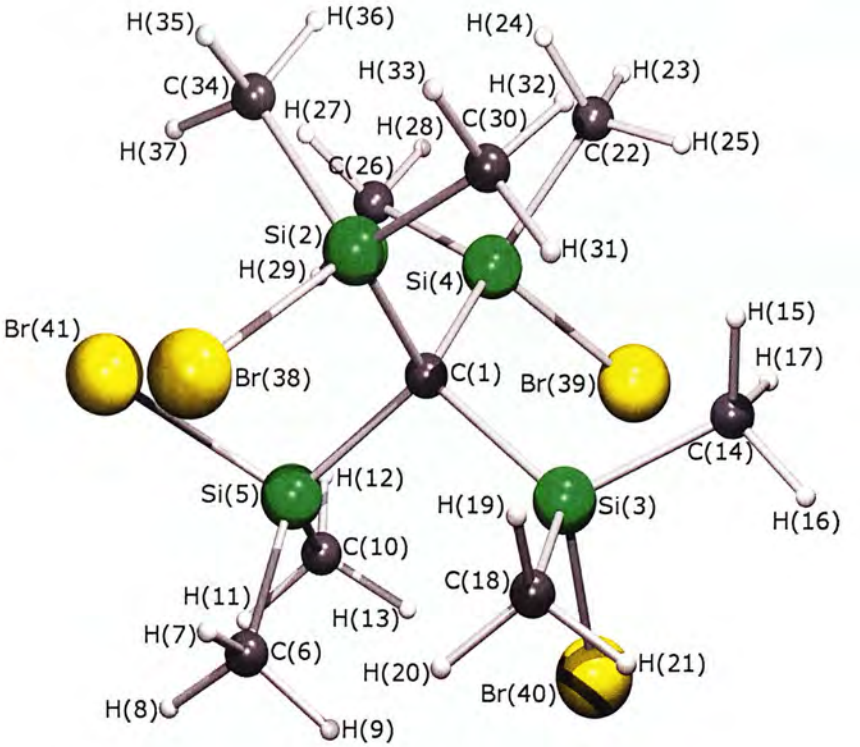


Conformer 3 (C_1 symmetry)

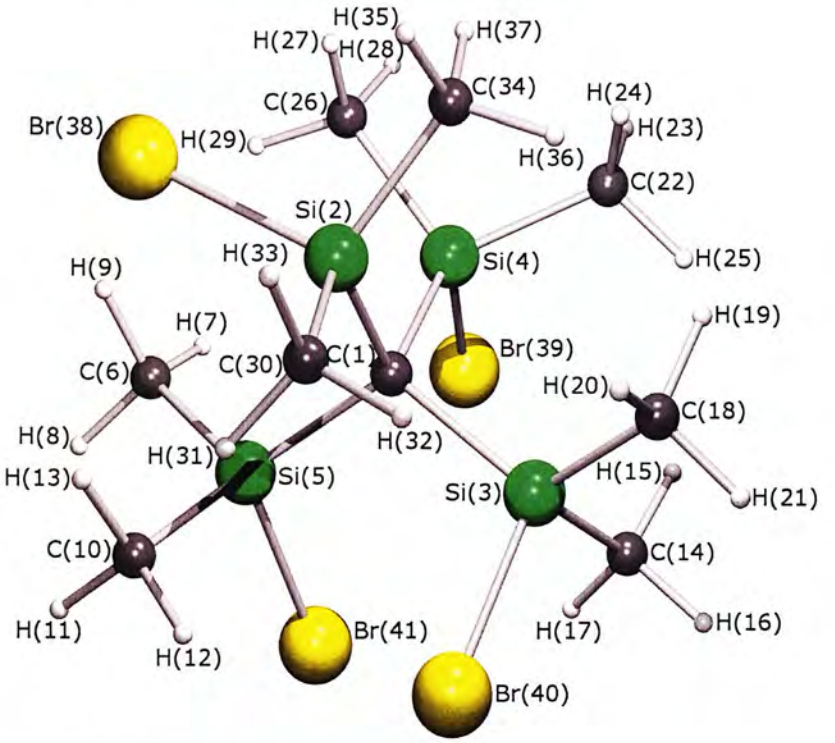
Figure 6.2. Gas-phase molecular structures of conformers 1, 2 and 3 of $C(SiMe_2F)_4$.



Conformer 1 (C_2 symmetry)



Conformer 2 (C_1 symmetry)



Conformer 3 (C_1 symmetry)

Figure 6.3. Gas-phase molecular structures of conformers 1, 2 and 3 of $C(SiMe_2Br)_4$.

For each conformer of each molecule, the amplitudes of vibration (u_{h1}) and curvilinear vibrational correction factors (k_{h1}) were then derived as detailed in Chapter 2.

The TINKER molecular mechanics package used in the SEMTEX refinement process does not have force constants for Si-Cl distances. However, since only the differences between peripheral-atom parameters are used in the refinement, it was decided that the silicon atoms could be replaced by carbon atoms for the purposes of the MM calculation. The values obtained from these calculations are immediately modified by the calculated difference set in order to reflect the MP2 calculation (which contains the silicon atoms), and so the final refined structure from SEMTEX is unaffected by this substitution.

6.2.2. Electron diffraction data

The GED data used in the refinements were previously collected at the University of Edinburgh by the procedure outlined in Chapter 2, using samples synthesised and characterised at Imperial College[5] using established methods[1].

6.2.3. Electron diffraction model

To describe the structure of $C(SiMe_2X)_4$ [X = Cl, F, Br], a three-conformer model with 43 geometric parameters was used in each case, comprising four bond lengths, nine bond angles and thirty torsion angles, as indicated by the *ab initio* calculations described above. These can be found in Tables 6.1 - 6.3. Two non-geometric parameters were also used to control the amounts of each conformer present in the mixture.

The equivalent bond lengths for the three conformers were almost identical when calculated at the MP2/6-311G* level of theory, and so it was determined that average parameters could be used for all three conformers with fixed differences between the conformers. The core-atom bond lengths were thus described by an average C-Si distance, *r*C-Si av. (p_1) and a difference *r*C-Si diff. ($p_{1\text{-}2}$), defined as the difference between the average C(1)-Si distance and the average Si-Me distance. *Ab initio*

calculations also showed that while in principle there are three different Si-C distances in a $SiMe_2X$ group, the differences are very small, and therefore these were also fixed to calculated values.

For the heavy-atom bond-angle parameters, an average parameter was deemed sufficient to describe the position of the methyl groups relative to the central carbon atom, ∠C-Si-C av. (p_3), with the differences between conformers fixed (as for the Si-C bond lengths) to values calculated *ab initio*. A parameter to describe the positions of the methyl-group carbons within a $SiMe_2X$ group relative to each other was also included, ∠Me-Si-Me av. (p_4). To describe the central Si-C-Si angles, more parameters were required. An overall 3-conformer average was used, ∠Si-C-Si av. (p_5), along with a difference between the average for conformer 1 and the average for conformers 2 and 3, ∠Si-C-Si diff. [∠Si-C-Si av. – ∠Si-C-Si av.(c1)] (p_6) Furthermore, an individual difference parameter was used for each conformer, e.g. ∠Si-C-Si diff.(c1) [∠Si-C-Si av.(c1) – ∠Si(2)-C(1)-Si(4)(c1)] and the analogous parameter for each of the other two conformers (p_{7-9}), to describe the largest difference between angles in each case. The further small differences between individual angles were fixed to values calculated *ab initio*.

Twenty-two parameters were also included to describe the starting positions of the hydrogen atoms. These comprised *r*C-H (p_{10}), ∠Si-C-H (p_{11}), and twenty parameters to describe the torsions of the methyl groups in the three different conformers about their respective Si-C bonds: four for conformer 1 (p_{12-15}), eight for conformer 2 (p_{16-23}), and eight for conformer 3 (p_{24-31}). In the SEMTEX refinements, p_{10} and p_{11} represented the mean values of those for all hydrogen atoms in the molecule. The specific torsions described by each parameter are given in Table 5.4.

Finally, twelve parameters were used to describe the positions of the halogen atoms. These were the average distance *r*Si-X (p_{32}), average angle ∠C-Si-X (p_{33}) and ten torsional angles to describe the position of the different *tert*-butyl groups around the Si-C(1) bond, two for conformer 1 (p_{34-35}), four for conformer 2 (p_{36-39}) and four for conformer 3 (p_{40-43}). In the SEMTEX refinements, p_{32} and p_{33} represented the mean values of those for all halogen atoms in the molecule. The specific torsions described by individual parameters can again be found in Table 6.4.

Table 6.1. Refined and calculated parameters for $C(SiMe_2Cl)_4$ (distances in pm, angles in °) from the SARACEN and SEMTEX studies.

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_1	*r*Si-C av.	189.0	190.8(2)	191.0(2)	
p_2	*r*Si-C diff.	0.5	-0.6(11)	-0.4(7)	
p_3	∠C-Si-Me av.	115.0	117.6(6)	113.3(5)	
p_4	∠Me-Si-Me av.	107.3	105.7(4)	103.1(5)	
p_5	∠Si-C-Si av.	109.4	111.8(9)	109.8(7)	
p_6	∠Si-C-Si diff.	5.7	6.2(15)	6.4(9)	5.7(15)
p_7	∠Si-C-Si diff. c1	-1.8	-3.6(11)	-3.0(9)	
p_8	∠Si-C-Si diff. c2	-2.3	-2.0(10)	-2.5(10)	-2.3(10)
p_9	∠Si-C-Si diff. c3	-3.2	-3.1(7)	-3.2(7)	-3.2(10)
p_{10}	*r*C-H av.	109.0	107.2(6)	107.8(5)	
p_{11}	∠Si-C-H av.	110.7	111.1(11)	110.9(9)	
p_{12}	Metor1 c1	66.4	65.0(51)	66.5(49)	66.4(50)
p_{13}	Metor2 c1	178.7	178.9(51)	177.4(49)	178.7(50)
p_{14}	Metor3 c1	-55.1	-57.4(50)	-57.7(47)	-55.1(50)
p_{15}	Metor4 c1	177.9	178.5(51)	176.9(49)	177.9(50)
p_{16}	Metor1 c2	-60.6	-60.6(52)	-60.7(49)	-60.6(50)
p_{17}	Metor2 c2	-67.3	-67.5(52)	-67.6(49)	-67.3(50)
p_{18}	Metor3 c2	-179.3	-179.1(52)	-179.6(50)	-179.3(50)

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{19}	Metor4 c2	55.5	55.3(52)	55.2(49)	55.5(50)
p_{20}	Metor5 c2	-178.8	-178.9(52)	-179.0(49)	-178.8(50)
p_{21}	Metor6 c2	-68.9	-68.9(37)	-69.0(35)	-68.9(50)
p_{22}	Metor7 c2	-179.2	-179.0(52)	-179.3(50)	-179.2(50)
p_{23}	Metor8 c2	-66.9	-67.1(52)	-67.1(49)	-66.9(50)
p_{24}	Metor1 c3	-59.6	-59.5(52)	-59.4(50)	-59.6(50)
p_{25}	Metor2 c3	-62.6	-62.5(52)	-62.5(50)	-62.6(50)
p_{26}	Metor3 c3	-177.7	-177.5(52)	-177.3(50)	-177.7(50)
p_{27}	Metor4 c3	60.6	60.8(52)	60.4(50)	60.6(50)
p_{28}	Metor5 c3	-177.6	-177.4(52)	-177.3(50)	-177.6(50)
p_{29}	Metor6 c3	-63.5	-63.7(52)	-63.6(50)	-63.5(50)
p_{30}	Metor7 c3	-176.1	-176.1(52)	-175.7(50)	-176.1(50)
p_{31}	Metor8 c3	-63.8	-64.0(52)	-63.5(50)	-63.8(50)
p_{32}	*r*Si-Cl av.	209.7	208.6(2)	208.5(2)	
p_{33}	∠C-Si-Cl av.	109.6	110.1(5)	110.8(5)	
p_{34}	$SiMe_2$Cltor1 c1	75.0	70.8(9)	83.3(11)	
p_{35}	$SiMe_2$Cltor2 c1	-40.3	-18.2(7)	-18.0(9)	
p_{36}	$SiMe_2$Cltor1 c2	162.0	161.5(50)	162.6(49)	162.0(50)
p_{37}	$SiMe_2$Cltor2 c2	159.9	161.4(49)	161.9(49)	159.9(50)

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{38}	$SiMe_2$Cltor3 c2	162.3	162.8(49)	163.3(49)	162.3(50)
p_{39}	$SiMe_2$Cltor4 c2	38.7	41.5(49)	40.8(48)	38.7(50)
p_{40}	$SiMe_2$Cltor1 c3	-165.9	-163.9(49)	-165.2(43)	-165.9(50)
p_{41}	$SiMe_2$Cltor2 c3	-35.2	-35.9(49)	-35.3(42)	-35.2(50)
p_{42}	$SiMe_2$Cltor3 c3	77.3	77.2(49)	77.0(45)	77.3(50)
p_{43}	$SiMe_2$Cltor4 c3	-43.2	-43.8(50)	-46.9(44)	-43.2(50)
Dependent parameters					
rC(1)-Si(2) conf 1		189.8	190.9	191.3	
rC(1)-Si(3) conf 1		188.9	190.1	190.3	
rC(1)-Si(2) conf 2		188.9	190.2	190.5	
rC(1)-Si(3) conf 2		189.6	190.6	191.2	
rC(1)-Si(4) conf 2		189.4	190.6	190.9	
rC(1)-Si(5) conf 2		189.2	190.4	190.6	
rC(1)-Si(2) conf 3		188.8	190.2	190.4	
rC(1)-Si(3) conf 3		189.8	190.6	191.3	
rC(1)-Si(4) conf 3		189.4	190.7	190.8	
rC(1)-Si(5) conf 3		189.1	190.3	190.6	
rC(methyl)-Si average		188.8	191.1	191.2	
∠Si(2)-C(1)-Si(4) confl		111.4	113.1	111.5	

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
	∠Si(3)-C(1)-Si(5) conf1	113.2	116.7	114.5	
	∠Si(2)-C(1)-Si(4) conf2	105.3	107.7	105.3	
	∠Si(3)-C(1)-Si(5) conf2	107.7	109.7	107.9	
	∠Si(2)-C(1)-Si(4) conf3	108.3	110.7	108.7	
	∠Si(3)-C(1)-Si(5) conf3	111.5	113.9	111.9	
	Composition parameters				
f_1	fraction of conf. 1	70.5	74.5	75.0	
f_2	fraction of conf. 2	15.5	12.5	8.0	
f_3	fraction of conf. 3	14.0	13.0	17.0	

Table 6.2. Refined and calculated parameters for $C(SiMe_2F)_4$ (distances in pm, angles in °) from the SARACEN and SEMTEX studies.

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_1	rSi-C av.	187.5	187.2(1)	187.4(1)	
p_2	rSi-C diff.	3.0	0.9(4)	-0.4(4)	
p_3	∠C-Si-Me av.	114.5	116.1(3)	116.0(2)	
p_4	∠Me-Si-Me av.	108.8	107.4(4)	106.9(2)	
p_5	∠Si-C-Si av.	109.9	107.5(4)	108.0(3)	
p_6	∠Si-C-Si diff.	3.5	2.8(7)	3.5(6)	3.5
p_7	∠Si-C-Si diff. c1	0.4	0.5(7)	0.9(4)	0.4
p_8	∠Si-C-Si diff. c2	3.5	3.4(7)	3.8(7)	3.5
p_9	∠Si-C-Si diff. c3	2.7	2.6(7)	2.4(7)	2.7
p_{10}	rC-H av.	109.2	109.6(2)	109.3(2)	
p_{11}	∠Si-C-H av.	112.0	110.2(6)	109.8(5)	
p_{12}	Metor1 c1	-59.2	-56.7(48)	-59.4(47)	-59.2
p_{13}	Metor2 c1	178.7	179.7(48)	182.6(45)	178.7
p_{14}	Metor3 c1	-55.1	-51.2(47)	-54.5(45)	-55.1
p_{15}	Metor4 c1	177.9	178.4(48)	178.3(47)	177.9
p_{16}	Metor1 c2	-50.9	-50.8(48)	-50.8(47)	-50.9
p_{17}	Metor2 c2	-74.6	-74.4(48)	-74.9(47)	-74.6
p_{18}	Metor3 c2	-175.7	-175.5(48)	-175.9(47)	-175.7

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{19}	Metor4 c2	41.0	41.0(48)	40.5(47)	41.0
p_{20}	Metor5 c2	-170.7	-170.7(48)	-170.9(47)	-170.7
p_{21}	Metor6 c2	-72.8	-72.8(48)	-73.2(47)	-72.8
p_{22}	Metor7 c2	-169.6	-169.6(48)	-169.6(47)	-169.6
p_{23}	Metor8 c2	52.6	52.5(48)	52.3(47)	52.6
p_{24}	Metor1 c3	-17.0	-17.0(48)	-16.9(47)	-17.0
p_{25}	Metor2 c3	-62.8	-62.9(48)	-62.3(46)	-62.8
p_{26}	Metor3 c3	-170.4	-170.5(48)	-170.6(47)	-170.4
p_{27}	Metor4 c3	58.1	58.1(48)	58.4(47)	58.1
p_{28}	Metor5 c3	-159.1	-159.1(48)	-159.0(47)	-159.1
p_{29}	Metor6 c3	-62.8	-63.0(48)	-62.6(47)	-62.8
p_{30}	Metor7 c3	-174.2	-174.4(48)	-174.0(47)	-174.2
p_{31}	Metor8 c3	-56.8	-57.0(48)	-56.7(47)	-56.8
p_{32}	rSi-F av.	162.5	160.5(1)	160.0(1)	
p_{33}	∠C-Si-F av.	105.7	105.4(6)	106.2(3)	
p_{34}	$SiMe_2$Ftor1 c1	43.6	48.9(29)	63.4(7)	
p_{35}	$SiMe_2$Ftor2 c1	-76.8	-82.0(16)	-80.3(7)	
p_{36}	$SiMe_2$Ftor1 c2	163.2	164.6(48)	165.9(40)	163.2
p_{37}	$SiMe_2$Ftor2 c2	31.5	27.6(47)	25.3(37)	31.5

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{38}	$SiMe_2$Ftor3 c2	165.4	165.9(48)	167.6(45)	165.4
p_{39}	$SiMe_2$Ftor4 c2	166.5	165.8(47)	169.9(45)	166.5
p_{40}	$SiMe_2$Ftor1 c3	-160.9	-162.7(46)	-159.8(43)	-160.9
p_{41}	$SiMe_2$Ftor2 c3	-46.6	-47.4(47)	-50.0(43)	-46.6
p_{42}	$SiMe_2$Ftor3 c3	79.7	80.8(48)	75.4(43)	79.7
p_{43}	$SiMe_2$Ftor4 c3	-36.5	-35.6(47)	-33.6(45)	-36.5
Dependent parameters					
rC(1)-Si(2) conf 1		190.0	188.3	187.6	
rC(1)-Si(3) conf 1		189.0	187.3	186.7	
rC(1)-Si(2) conf 2		189.2	187.5	186.7	
rC(1)-Si(3) conf 2		189.8	188.1	187.4	
rC(1)-Si(4) conf 2		189.6	187.9	187.1	
rC(1)-Si(5) conf 2		189.5	187.8	187.0	
rC(1)-Si(2) conf 3		189.1	187.4	186.6	
rC(1)-Si(3) conf 3		189.8	188.1	187.4	
rC(1)-Si(4) conf 3		189.6	187.9	187.0	
rC(1)-Si(5) conf 3		189.4	187.7	186.9	
rC(methyl)-Si average		186.5	186.9	187.5	
∠Si(2)-C(1)-Si(4) conf1		111.9	109.2	110.3	

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
	∠Si(3)-C(1)-Si(5) conf1	111.5	108.6	109.3	
	∠Si(2)-C(1)-Si(4) conf2	109.9	107.8	108.1	
	∠Si(3)-C(1)-Si(5) conf2	106.3	104.4	104.3	
	∠Si(2)-C(1)-Si(4) conf3	111.8	109.3	109.7	
	∠Si(3)-C(1)-Si(5) conf3	109.9	106.7	107.3	
	Composition parameters				
f_1	fraction of conf. 1	23.0	26.0	25.0	
f_2	fraction of conf. 2	48.3	41.0	42.0	
f_3	fraction of conf. 3	28.7	33.0	33.0	

Table 6.3. Refined and calculated parameters for $C(SiMe_2Br)_4$ (distances in pm, angles in °) from the SARACEN and SEMTEX studies. A LanL2DZ pseudopotential was used on the bromine atoms in the MP2/6-311G* calculations.

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_1	rSi-C av.	188.0	187.5(1)	187.8(2)	
p_2	rSi-C diff.	5.0	7.7(7)	8.6(6)	
p_3	aC-Si-Me av.	114.9	116.4(4)	114.8(4)	
p_4	aMe-Si-Me av.	103.5	103.5(6)	105.0(8)	
p_5	aSi-C-Si av.	110.1	111.5(9)	111.3(8)	
p_6	aSi-C-Si diff.	4.8	4.3(9)	4.4(5)	4.8(10)
p_7	Si-C-Si diff. c1	2.4	2.0(7)	2.3(6)	2.4(8)
p_8	Si-C-Si diff. c2	4.1	3.9(7)	4.0(7)	4.1(8)
p_9	Si-C-Si diff. c3	5.4	5.3(7)	5.3(7)	5.4(8)
p_{10}	rC-H av.	109.0	109.5(3)	109.8(3)	
p_{11}	aSi-C-H av.	110.7	111.6(8)	111.9(9)	
p_{12}	Metor1 c1	-59.2	-60.1(49)	-59.6(48)	-59.2(50)
p_{13}	Metor2 c1	-66.8	-69.6(49)	-68.3(48)	-66.8(50)
p_{14}	Metor3 c1	61.7	63.1(49)	61.3(48)	61.7(50)
p_{15}	Metor4 c1	178.1	178.0(49)	177.7(48)	178.1(50)
p_{16}	Metor1 c2	-60.4	-60.4(49)	-60.4(49)	-60.4(50)
p_{17}	Metor2 c2	-61.3	-61.3(49)	-61.5(49)	-61.3(50)
p_{18}	Metor3 c2	177.2	177.1(49)	177.1(49)	177.2(50)

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{19}	Metor4 c2	60.9	61.4(49)	60.8(49)	60.9(50)
p_{20}	Metor5 c2	179.4	179.5(49)	179.3(49)	179.4(50)
p_{21}	Metor6 c2	-66.5	-66.8(49)	-67.0(49)	-66.5(50)
p_{22}	Metor7 c2	177.3	177.3(49)	177.3(49)	177.3(50)
p_{23}	Metor8 c2	-62.0	-61.6(49)	-62.0(49)	-62.0(50)
p_{24}	Metor1 c3	-61.0	-61.1(49)	-60.9(49)	-61.0(50)
p_{25}	Metor2 c3	-61.6	-61.7(49)	-61.8(49)	-61.6(50)
p_{26}	Metor3 c3	-178.1	-178.2(49)	-178.0(49)	-178.1(50)
p_{27}	Metor4 c3	62.6	62.9(49)	62.6(49)	62.6(50)
p_{28}	Metor5 c3	-179.1	-179.2(49)	-179.0(49)	-179.1(50)
p_{29}	Metor6 c3	-62.5	-63.1(49)	-62.7(49)	-62.5(50)
p_{30}	Metor7 c3	-177.6	-177.7(49)	-177.5(49)	-177.6(50)
p_{31}	Metor8 c3	-61.6	-61.5(49)	-61.5(49)	-61.6(50)
p_{32}	rSi-Br av.	225.0	226.5(1)	226.3(1)	
p_{33}	aC-Si-Br av.	109.5	108.7(3)	108.8(3)	
p_{34}	$SiMe_2$Brtor1 c1	39.4	48.5(15)	54.1(12)	
p_{35}	$SiMe_2$Brtor2 c1	-74.5	-80.3(22)	-78.0(13)	
p_{36}	$SiMe_2$Brtor1 c2	161.2	163.0(39)	162.3(38)	161.2(50)
p_{37}	$SiMe_2$Brtor2 c2	39.3	40.9(29)	42.4(35)	39.3(50)

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
p_{38}	$SiMe_2$Brtor3 c2	161.3	161.5(41)	161.6(35)	161.3(50)
p_{39}	$SiMe_2$Brtor4 c2	158.6	157.4(20)	160.6(27)	158.6(50)
p_{40}	$SiMe_2$Brtor1 c3	-166.7	-170.2(38)	-169.5(41)	-166.7(50)
p_{41}	$SiMe_2$Brtor2 c3	-35.8	-36.0(42)	-36.0(31)	-35.8(50)
p_{42}	$SiMe_2$Brtor3 c3	78.0	75.0(43)	75.6(41)	78.0(50)
p_{43}	$SiMe_2$Brtor4 c3	-42.7	-35.4(41)	-37.3(37)	-42.7(50)
Dependent parameters					
rC(1)-Si(2) conf 1		191.8	191.9	192.5	
rC(1)-Si(3) conf 1		190.9	191.0	191.5	
rC(1)-Si(2) conf 2		190.9	191.1	191.7	
rC(1)-Si(3) conf 2		191.6	191.7	192.3	
rC(1)-Si(4) conf 2		191.4	191.6	192.1	
rC(1)-Si(5) conf 2		191.3	191.5	192.0	
rC(1)-Si(2) conf 3		190.8	191.1	191.7	
rC(1)-Si(3) conf 3		191.6	191.8	192.4	
rC(1)-Si(4) conf 3		191.5	191.4	192.0	
rC(1)-Si(5) conf 3		191.2	191.4	192.0	
rC(methyl)-Si average		186.3	184.1	184.2	
∠Si(2)-C(1)-Si(4) conf1		113.7	114.7	114.7	

no.	Parameter	MP2/ 6-311G* (r_e)	SARACEN (r_{h1})	SEMTEX (r_{h1})	Restraint
	∠Si(3)-C(1)-Si(5) conf1	111.5	112.7	112.3	
	∠Si(2)-C(1)-Si(4) conf2	109.8	111.3	111.1	
	∠Si(3)-C(1)-Si(5) conf2	105.6	107.3	107.1	
	∠Si(2)-C(1)-Si(4) conf3	113.3	114.7	114.5	
	∠Si(3)-C(1)-Si(5) conf3	107.9	109.3	109.1	
	Composition parameters				
f_1	fraction of conf. 1	73.1	57.5	57.0	
f_2	fraction of conf. 2	13.7	22.0	22.0	
f_3	fraction of conf. 3	13.2	20.5	21.0	

Table 6.4. Specific torsions described by each parameter in the electron diffraction model. Parameters for conformers 2 and 3 describe the same torsion in all cases.

Parameter	Specific torsion
Conformer 1	
Metor1	C(1)-Si(4)-C(22)-H(23)
Metor2	C(1)-Si(4)-C(26)-H(27)
Metor3	C(1)-Si(3)-C(14)-H(15)
Metor4	C(1)-Si(3)-C(18)-H(19)
SiMeCltor1	Si(2)-C(1)-Si(4)-Cl(39)

Parameter	Specific torsion
SiMeCltor2	Si(5)-C(1)-Si(3)-Cl(40)
Conformers 2 and 3	
Metor1	C(1)-Si(4)-C(26)-H(29)
Metor2	C(1)-Si(4)-C(22)-H(24)
Metor3	C(1)-Si(2)-C(30)-H(33)
Metor4	C(1)-Si(2)-C(34)-H(36)
Metor5	C(1)-Si(3)-C(14)-H(16)
Metor6	C(1)-Si(3)-C(18)-H(19)
Metor7	C(1)-Si(5)-C(10)-H(11)
Metor8	C(1)-Si(5)-C(6)-H(9)
SiMeCltor1	Si(2)-C(1)-Si(4)-Cl(39)
SiMeCltor2	Si(4)-C(1)-Si(2)-Cl(38)
SiMeCltor3	Si(5)-C(1)-Si(3)-Cl(40)
SiMeCltor4	Si(3)-C(1)-Si(5)-Cl(41)

6.3. Results

6.3.1. Tetrakis(-chlorodimethylsilyl)methane

6.3.1.1. Theoretical methods

The structure of $C(SiMe_2Cl)_4$ was determined *ab initio*. A potential-energy surface search involving rotation of the $SiMe_2Cl$ groups around the Si-C bonds found three potential conformers that could be present in significant quantity in the experimental mixture: conformer 1, exhibiting C_2 symmetry, and conformers 2 and 3, each exhibiting C_1 symmetry. The molecular geometry of each conformer of $C(SiMe_2Cl)_4$ at the MP2/6-311G* level can be found in Table 6.1.

The Gibbs free energies of the conformers were calculated at the Hartree-Fock level of theory with the 6-31G* basis set. The energy differences between the three conformers calculated at this level are given in Table 6.5, along with the relative abundances calculated using the Boltzmann distribution.

Table 6.5. Total energies and energy differences between conformers calculated at the HF/6-31G* level. Energy differences are relative to conformer 1, the lowest-energy conformer. Conformers marked with a * were not considered to be present in sufficient quantity to be included in the refinement model.

Conformer	Symmetry	Total energy / kJ mol^{-1}	Energy difference / kJ mol^{-1}	Abundance / %
1	C_2	-8791313.94	0.0	67.8
2	C_1	-8791305.71	+8.2	14.9
3	C_1	-8791305.31	+8.6	13.3
4*	D_2	-8791303.42	+10.6	2.0
5*	C_2	-8791293.04	+20.9	0.2
6*	C_1	-8791295.37	+18.6	0.9
7*	C_1	-8791295.36	+18.6	0.9

From these energy differences, therefore, it is probable that only three of these conformers would be present in the experimental mixture in significant quantity, and so three conformers were included in the model written for the least-squares refinement of the GED data. The predicted amounts for each conformer were recalculated on the basis that only three were present, which predicted a 70.5 : 15.5 : 14.0 mixture, as shown in Table 6.1.

At the outset of the SEMTEX refinement process, the heavy-atom positions were fixed as calculated using SARACEN, and both MP2/6-311G* and MM3 calculations were performed on the peripheral atoms. Table 6.6 shows the resulting peripheral-atom parameters from both the MM and MP2 calculations for one Me_2Cl group of each conformer. The full set of peripheral-atom parameters is given in Table S6a.1 in the Supplementary Information.

Table 6.6. Peripheral-atom parameters for $C(SiMe_2Cl)_4$ calculated with the MP2/6-311G* and MM3 methods. All bond lengths are in pm and angles in °. Internuclear distances are the calculated (r_e) values. In the MM3 calculations, Si atoms were replaced by C atoms; however, only the *differences* between the calculated parameters are relevant to the GED refinements.

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H	
Conformer 1	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	108.0	*109.9*	119.5	*113.2*	55.4	*80.3*
H(8)-C(6)-Si(5)-C(1)	109.8	*111.0*	102.5	*108.5*	174.6	*-162.6*
H(9)-C(6)-Si(5)-C(1)	109.0	*110.7*	111.0	*109.6*	-68.9	*-46.6*
H(11)-C(10)-Si(5)-C(1)	109.7	*111.0*	106.7	*109.0*	-155.5	*-152.4*
H(12)-C(10)-Si(5)-C(1)	109.1	*110.8*	117.7	*109.8*	-36.8	*-34.0*
H(13)-C(10)-Si(5)-C(1)	108.6	*110.7*	108.2	*110.8*	85.7	*89.2*

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H	
Conformer 2						
H(7)-C(6)-Si(5)-C(1)	109.2	*110.8*	115.9	*109.8*	62.6	*62.1*
H(8)-C(6)-Si(5)-C(1)	109.4	*111.0*	106.7	*109.4*	176.9	*-178.1*
H(9)-C(6)-Si(5)-C(1)	109.1	*110.7*	109.8	*109.9*	-60.2	*-58.5*
H(11)-C(10)-Si(5)-C(1)	109.4	*111.0*	106.4	*109.5*	180.0	*-174.0*
H(12)-C(10)-Si(5)-C(1)	109.2	*110.8*	116.0	*110.1*	-60.7	*-54.1*
H(13)-C(10)-Si(5)-C(1)	109.1	*110.7*	109.9	*110.0*	63.2	*66.8*
Conformer 3						
H(7)-C(6)-Si(5)-C(1)	109.2	*110.9*	116.2	*109.5*	57.5	*65.0*
H(8)-C(6)-Si(5)-C(1)	109.5	*111.0*	109.1	*109.3*	176.9	*-176.0*
H(9)-C(6)-Si(5)-C(1)	108.8	*110.5*	106.6	*111.0*	-66.3	*-56.2*
H(11)-C(10)-Si(5)-C(1)	109.5	*111.0*	106.4	*109.3*	178.2	*177.9*
H(12)-C(10)-Si(5)-C(1)	109.1	*110.8*	115.4	*110.5*	-62.2	*-62.4*
H(13)-C(10)-Si(5)-C(1)	109.2	*110.8*	110.4	*109.8*	61.9	*58.7*

Parameter	*r*Si-Cl		∠C-Si-Cl	
Conformer 1				
Cl(39)-Si(4)-C(1)	207.2	*178.8*	114.2	*115.8*
Cl(40)-Si(5)-C(1)	207.0	*179.0*	115.6	*115.6*
Conformer 2				
Cl(38)-Si(2)-C(1)	206.0	*178.7*	111.5	*113.7*

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H
Cl(39)-Si(4)-C(1)	206.0	*178.8*	111.4	*113.1*	
Cl(40)-Si(5)-C(1)	206.8	*178.6*	111.7	*112.8*	
Cl(41)-Si(3)-C(1)	207.2	*178.9*	115.5	*114.3*	
Conformer 3					
Cl(38)-Si(2)-C(1)	206.9	*178.5*	115.2	*114.5*	
Cl(39)-Si(4)-C(1)	206.3	*178.8*	113.8	*114.5*	
Cl(40)-Si(5)-C(1)	206.5	*178.7*	112.9	*113.8*	
Cl(41)-Si(3)-C(1)	206.6	*179.0*	111.2	*112.9*	

6.3.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311G* level. Forty-three geometric parameters were refined along with nine groups of vibrational amplitudes. Two parameters to describe the respective amounts of the three conformers in the experimental mixture were also included. Thirty-two geometric and four amplitude restraints were applied according to the SARACEN method, with most of the geometric restraints applying to the torsional parameters for the two less-abundant C_1 conformers. These can be found in Table 6.1. The final R factors for the refinement were $R_G = 0.106$ and $R_D = 0.070$. Interatomic distances and corresponding amplitudes of vibration are given in Table S6a.2 in the supplementary information. Final experimental coordinates from the SARACEN GED analysis are given in Table S6a.3, with the correlation matrix in Table S6a.4.

6.3.3. SEMTEX refinement

The starting parameters were as for the SARACEN refinement. Once all forty-three geometric parameters and nine groups of vibrational amplitudes were refined according to the SARACEN method, the SEMTEX code was activated and the peripheral-atom positions updated computationally.

As for the SARACEN refinement, all forty-three geometric parameters were refined along with nine groups of vibrational amplitudes. Thirty-two geometric and four amplitude restraints were applied using the SARACEN method. After refinement, the conformational composition was determined. First, the fraction parameter describing the amount of conformer 1 in the sample was altered in order to find which value gave the lowest R factor, with conformers 2 and 3 fixed at equal amounts. The resulting plot of R_G vs $R_{Gmin.}$ is shown in Figure 6.4. As can be seen, the R factor increases either side of the minimum at 75.0%. The 95% confidence limit[11] [where $R_G/R_G(min.) = 1.016$] shows that the uncertainty in this value is around ±4%. Once the fractional amount corresponding to the best fit was found for conformer 1, it was fixed and the amount of conformer 2 was altered to again find the minimum R factor, with the amount of conformer 3 defined by the amounts of the other two. The plot of R_G vs $R_{Gmin.}$ is shown in Figure 6.5. Again, the uncertainty can be seen to be around ±4%. From these, it can be seen that the best fit comes when the three conformers are present in a 75:8:17 ratio.

In the final refinement, the R factors were $R_G = 0.089$ and $R_D = 0.060$, obtained for a conformer ratio of 75.0 : 8.0 : 17.0 (conf 1 : conf 2 : conf 3). Figure 6.6 shows the molecular scattering intensities for the SEMTEX refinement, while Figure 6.7. shows the radial distribution curve. Table 6.1 lists the final refined parameters, and Table 6.7 shows the resulting peripheral-atom parameters for one Me_2Cl group of each conformer. The relative energies of the conformers calculated from this composition are given in Table 6.8. Interatomic distances and corresponding amplitudes of vibration are given in Table S6a.5 and final experimental coordinates from the SEMTEX GED analysis are given in Table S6a.6. The correlation matrix is given in Table S6a.7, and the full set of peripheral-atom parameters in Table S6a.8.

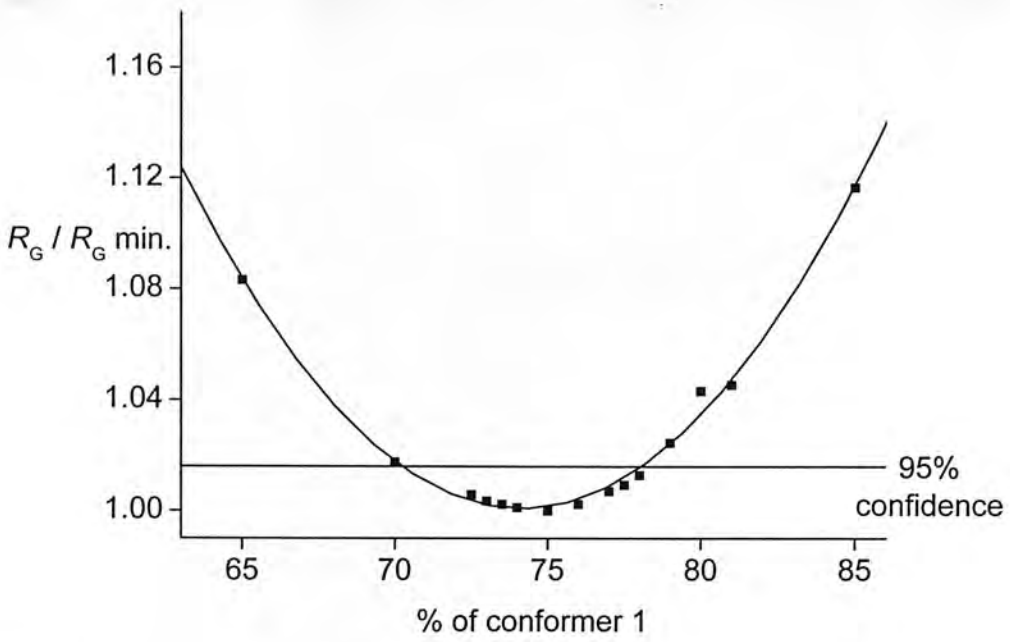


Figure 6.4. R_G / R_G min. for varying amounts of conformer 1 in the SEMTEX refinement of $C(SiMe_2Cl)_4$. Conformers 2 and 3 were fixed at equal amounts for each different amount of conformer 1.

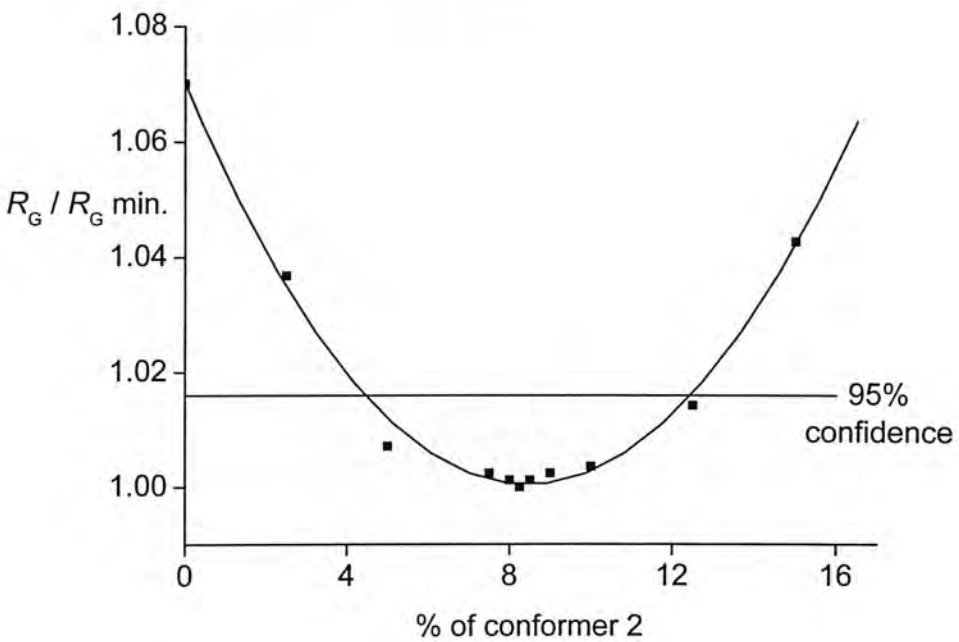


Figure 6.5. R_G / R_G min. for varying amounts of conformer 2 in the SEMTEX refinement of $C(SiMe_2Cl)_4$. Conformer 1 amount was fixed at 75% for each refinement, with the amount of conformer 3 defined by the amounts of the other two.

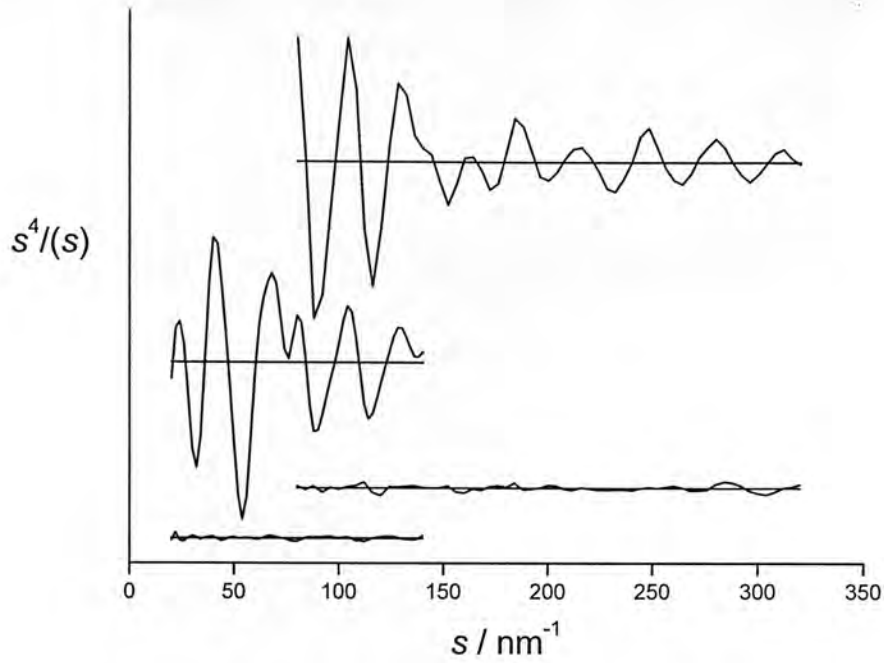


Figure 6.6. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $C(SiMe_2Cl)_4$.

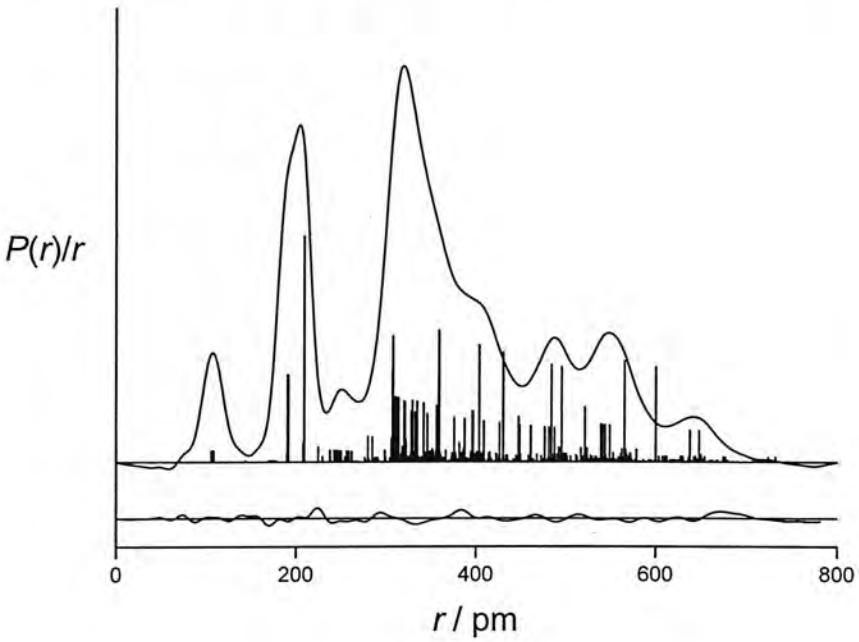


Figure 6.7. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $C(SiMe_2Cl)_4$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_{Si} - f_{Si})$.

Table 6.7. C-H bond lengths, Si-C-H bond angles and C-Si-C-H bond torsions for $C(SiMe_2Cl)_4$ calculated with the SEMTEX method. All bond lengths are in pm and angles in °. Interatomic distances are the refined (r_{h1}) values.

Parameter	*r*C-H	∠Si-C-H	∅C-Si-C-H
Conformer 1			
H(7)-C(6)-Si(5)-C(1)	106.6	119.5	61.0
H(8)-C(6)-Si(5)-C(1)	108.5	102.5	178.1
H(9)-C(6)-Si(5)-C(1)	107.6	111.0	-65.6
H(11)-C(10)-Si(5)-C(1)	108.3	106.7	-177.0
H(12)-C(10)-Si(5)-C(1)	107.8	117.8	-58.7
H(13)-C(10)-Si(5)-C(1)	107.2	108.3	64.6
Conformer 2			
H(7)-C(6)-Si(5)-C(1)	107.8	116.0	54.1
H(8)-C(6)-Si(5)-C(1)	108.0	106.7	173.9
H(9)-C(6)-Si(5)-C(1)	107.7	109.8	-66.6
H(11)-C(10)-Si(5)-C(1)	108.0	106.4	-179.6
H(12)-C(10)-Si(5)-C(1)	107.8	116.1	-60.2
H(13)-C(10)-Si(5)-C(1)	107.8	109.9	61.4
Conformer 3			
H(7)-C(6)-Si(5)-C(1)	107.9	116.2	57.2
H(8)-C(6)-Si(5)-C(1)	108.1	106.6	176.1
H(9)-C(6)-Si(5)-C(1)	107.4	109.2	-64.1
H(11)-C(10)-Si(5)-C(1)	108.1	106.5	-176.0

Parameter	*r*C-H	∠Si-C-H	ϕC-Si-C-H
H(12)-C(10)-Si(5)-C(1)	107.7	115.5	-56.3
H(13)-C(10)-Si(5)-C(1)	107.8	110.5	64.8
Parameter	*r*Si-Cl	∠C-Si-Cl	
Conformer 1			
Cl(39)-Si(4)-C(1)	208.8	113.2	
Cl(40)-Si(5)-C(1)	209.1	111.9	
Conformer 2			
Cl(38)-Si(2)-C(1)	207.8	109.3	
Cl(39)-Si(4)-C(1)	207.9	109.2	
Cl(40)-Si(5)-C(1)	208.7	109.5	
Cl(41)-Si(3)-C(1)	209.0	113.1	
Conformer 3			
Cl(38)-Si(2)-C(1)	208.7	112.8	
Cl(39)-Si(4)-C(1)	208.1	111.5	
Cl(40)-Si(5)-C(1)	208.3	110.6	
Cl(41)-Si(3)-C(1)	208.4	109.0	

Table 6.8. Energy differences and relative abundances of the three conformers from the MP2/6-311G* level of theory and from the SEMTEX refinement. The uncertainties for the energy differences are estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 1.

Conformer	Symmetry	Energy difference / kJ mol^{-1}		Abundance / %	
		HF/ 6-31G*	SEMTEX	HF/ 6-31G*	SEMTEX
1	C_2	0.0	0.0	70.5	75.0(20)
2	C_1	+8.2	+10.6(6)	15.5	8.0(28)
3	C_1	+8.6	+8.1(6)	14.0	17.0(28)

6.3.2. Tetrakis(fluorodimethylsilyl)methane

6.3.2.1. Theoretical methods

The structure of $C(SiMe_2F)_4$ was predicted *ab initio*. A potential-energy surface search involving rotation of the *tert*-butyl groups around the Si-C bonds found three potential conformers of $C(SiMe_2F)_4$ which could be present in significant quantity in the experimental mixture: conformer 1, exhibiting C_2 symmetry, and conformers 2 and 3, each exhibiting C_1 symmetry. The molecular geometry of each conformer of $C(SiMe_2F)_4$ at the MP2/6-311G* level can be found in Table 6.2.

The Gibbs free energies of the conformers were calculated at the Hartree-Fock level of theory with the 6-31G* basis set. The energy differences between the three conformers calculated at this level are given in Table 6.9, along with the relative abundances calculated using the Boltzmann distribution.

Table 6.9. Total energies and energy differences between conformers calculated at the HF/6-31G* level. Energy differences are relative to conformer 2, the lowest-energy conformer. Conformers marked with a * were not included in the refinement model.

Conformer	Symmetry	Total energy / kJ mol^{-1}	Energy difference / kJ mol^{-1}	Abundance / %
1	C_2	-5010977.848	+ 0.2	22.4
2	C_1	-5010978.003	0.00	46.8
3	C_1	-5010976.233	+ 1.8	27.8
4*	D_2	-5010969.756	+8.2	1.0
5*	C_2	-5010964.484	+13.5	0.4
6*	C_1	-5010964.132	+13.9	0.8
7*	C_1	-5010964.130	+13.9	0.8

From these energy differences, therefore, it is possible that all three of these conformers would be present in the experimental mixture, and so three conformers were included in the model written for the least-squares refinement of the GED data. The relative conformer amounts were recalculated based on a three-conformer mixture, which predicted relative abundances of 23.0 : 48.3 : 28.7. However, the energy differences between conformers were found to vary considerably on change of basis set, and so the relative abundances cannot be determined accurately from these calculations – the refinement might give very different amounts of each conformer.

At the outset of the SEMTEX refinement process, the heavy-atom positions were fixed as calculated using SARACEN, and both MP2/6-311G* and MM3 calculations were performed on the peripheral atoms. Table 6.10 shows the resulting peripheral-atom parameters from both the MM and MP2 calculations for one Me_2Cl group of each conformer. The full set of peripheral-atom parameters can be found in Table S6b.1 in the Supplementary Information.

Table 6.10. Peripheral-atom bond lengths, bond angles and bond torsions for $C(SiMe_2F)_4$ calculated with the MP2/6-311G* and MM3 methods. All bond lengths are in pm and angles in °. Internuclear distances are the calculated (r_e) values. In the MM3 calculations, Si atoms were replaced by C atoms However, only the *differences* between the calculated parameters are relevant to the GED refinements.

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H	
Conformer 1	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	109.1	*110.7*	112.8	*110.7*	53.5	*51.5*
H(8)-C(6)-Si(5)-C(1)	109.5	*111.0*	107.8	*109.4*	172.8	*171.2*
H(9)-C(6)-Si(5)-C(1)	109.3	*110.9*	112.6	*109.7*	-68.2	*-70.4*
H(11)-C(10)-Si(5)-C(1)	109.5	*111.0*	108.2	*109.9*	-133.7	*-161.6*
H(12)-C(10)-Si(5)-C(1)	108.9	*110.2*	116.7	*113.4*	-11.7	*-39.8*
H(13)-C(10)-Si(5)-C(1)	109.4	*110.9*	109.3	*108.7*	110.0	*80.2*
Conformer 2	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	109.5	*110.9*	106.6	*109.3*	172.2	*-163.7*
H(8)-C(6)-Si(5)-C(1)	109.1	*110.5*	113.2	*111.4*	-69.8	*-44.0*
H(9)-C(6)-Si(5)-C(1)	109.1	*110.8*	112.3	*110.8*	54.4	*77.5*
H(11)-C(10)-Si(5)-C(1)	109.5	*110.9*	107.6	*109.7*	-154.9	*157.2*
H(12)-C(10)-Si(5)-C(1)	109.2	*110.9*	114.9	*109.8*	-34.8	*-84.0*
H(13)-C(10)-Si(5)-C(1)	109.0	*110.4*	110.3	*112.8*	87.0	*35.7*
Conformer 3	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	109.1	*110.6*	112.6	*111.9*	56.2	*67.3*
H(8)-C(6)-Si(5)-C(1)	109.4	*111.0*	107.3	*109.4*	174.6	*-172.5*
H(9)-C(6)-Si(5)-C(1)	109.4	*110.8*	112.4	*109.6*	-67.1	*-53.9*

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H	
H(11)-C(10)-Si(5)-C(1)	109.4	*111.0*	107.2	*109.5*	-172.8	*-179.7*
H(12)-C(10)-Si(5)-C(1)	109.1	*110.6*	112.1	*111.6*	-54.3	*-59.5*
H(13)-C(10)-Si(5)-C(1)	109.3	*110.8*	112.6	*109.6*	69.1	*61.8*

Parameter	*r*Si-F		∠C-Si-F	
Conformer 1	MP2	MM3	MP2	MM3
F(39)-Si(4)-C(1)	164.3	*179.8*	105.0	*108.7*
F(41)-Si(5)-C(1)	164.9	*179.8*	103.8	*107.4*
Conformer 2	MP2	MM3	MP2	MM3
F(38)-Si(2)-C(1)	164.2	*179.9*	105.6	*107.1*
F(39)-Si(4)-C(1)	164.3	*180.0*	105.7	*107.4*
F(40)-Si(3)-C(1)	164.3	*180.2*	104.9	*112.1*
F(41)-Si(5)-C(1)	165.0	*180.1*	103.1	*106.8*
Conformer 3	MP2	MM3	MP2	MM3
F(38)-Si(2)-C(1)	164.7	*180.1*	104.5	*106.6*
F(39)-Si(4)-C(1)	164.2	*180.1*	105.6	*108.3*
F(40)-Si(3)-C(1)	164.2	*180.2*	104.2	*108.1*
F(41)-Si(5)-C(1)	164.1	*180.3*	104.5	*110.1*

6.3.2.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311G* level. Forty-three geometric parameters

were refined along with nine groups of vibrational amplitudes. Two parameters to describe the respective amounts of the three conformers in the experimental mixture were also included. Thirty-two geometric and four amplitude restraints were applied according to the SARACEN method, with most of the geometric restraints applying to the torsional parameters for the two less-abundant C_1 conformers. These can be found in Table 6.2. The final R factors for the refinement were found to be R_G = 0.071 and R_D = 0.076. Interatomic distances and corresponding amplitudes of vibration are given in Table S6b.2 in the supplementary information, and final experimental coordinates from the SARACEN GED analysis are given in Table S6b.3. The least-squares correlation matrix is given in Table S6b.4.

6.3.2.3. SEMTEX refinement

The starting parameters were as for the SARACEN refinement. Once all forty-three geometric parameters and nine groups of vibrational amplitudes were refined according to the SARACEN method, the SEMTEX code was activated and the peripheral-atom positions updated computationally. As previously detailed, carbon atoms had to be substituted for the silicon atoms for the MM calculations. However, the peripheral-atom positions returned at this stage are immediately modified by the difference set (MM – MP2) calculated for SEMTEX. As the MP2 calculation is done using the correct core atoms, frozen in space according to their coordinates from GED, the inaccuracy in the peripheral-atom positions returned by MM is negated and does not manifest itself in the final SEMTEX structure.

As for the SARACEN refinement, all forty-three geometric parameters were refined along with nine groups of vibrational amplitudes. Thirty-two geometric and four amplitude restraints were applied using the SARACEN method. After refinement, the conformational composition was determined. First, the fraction parameter describing the amount of conformer 2 in the sample was altered in order to find which value gave the lowest R factor, with conformers 1 and 3 fixed at equal amounts. The resulting plot of R_G v $R_{Gmin.}$ is shown in Figure 6.8. As can be seen, the R factor increases either side of the minimum at 42.0%. The 95% confidence limit[11] [where R_G/R_G(min.) = 1.016] shows that the uncertainty in this value (2 estimated

standard deviations) is around ±4%. Once the fractional amount corresponding to best fit was found for conformer 2, it was fixed and the amount of conformer 1 was altered to again find the minimum R factor, with the amount of conformer 3 defined by the amounts of the other two. The plot of R_G vs $R_{Gmin.}$ is shown in Figure 6.9 – this time, the uncertainty can be seen to be around ±5%. From these, it can be seen that the best fit comes when the three conformers are present in a 25:42:33 ratio.

In the final refinement, the R factors were $R_G = 0.059$ and $R_D = 0.048$, obtained for a conformer ratio of 25.0 : 42.0 : 33.0 (conf 1 : conf 2 : conf 3). Figure 6.10 shows the molecular scattering intensities for the SEMTEX refinement, while Figure 6.11 shows the radial distribution curve. Table 6.2 lists the final refined parameters, and Table 6.11 shows the resulting peripheral-atom parameters for one Me_2Cl group of each conformer. The relative energies of the conformers calculated from this composition are given in Table 6.12. Interatomic distances and corresponding amplitudes of vibration are given in Table S6b.5 and final experimental coordinates from the SEMTEX GED analysis are given in Table S6b.6. The least-squares correlation matrix is given in Table S6b.7. The full set of peripheral-atom parameters is given in Table S6b.8.

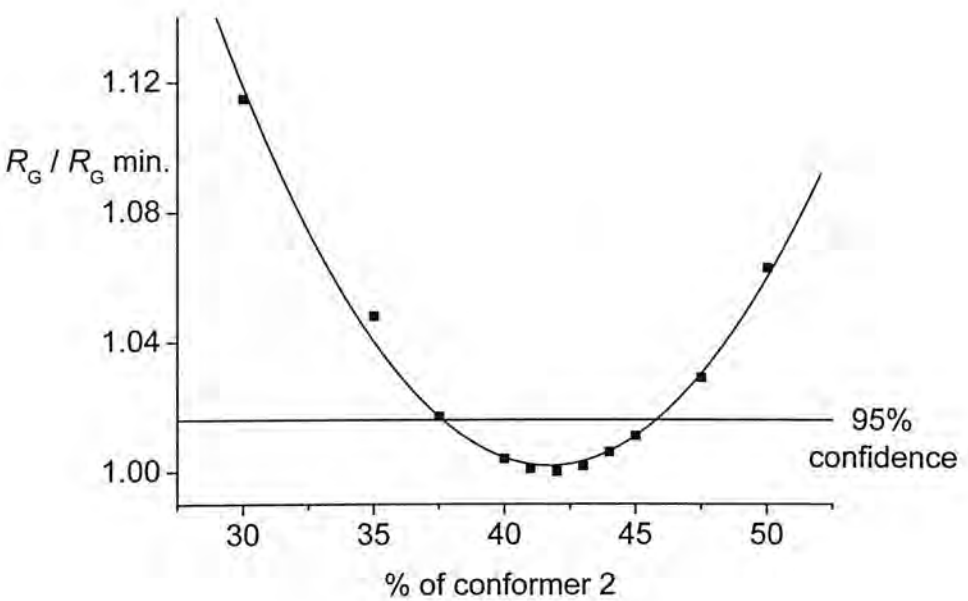


Figure 6.8. R_G / R_G min. for varying amounts of conformer 2 in the SEMTEX refinement of $C(SiMe_2F)_4$. Conformers 2 and 3 were fixed at equal amounts for each different amount of conformer 1.

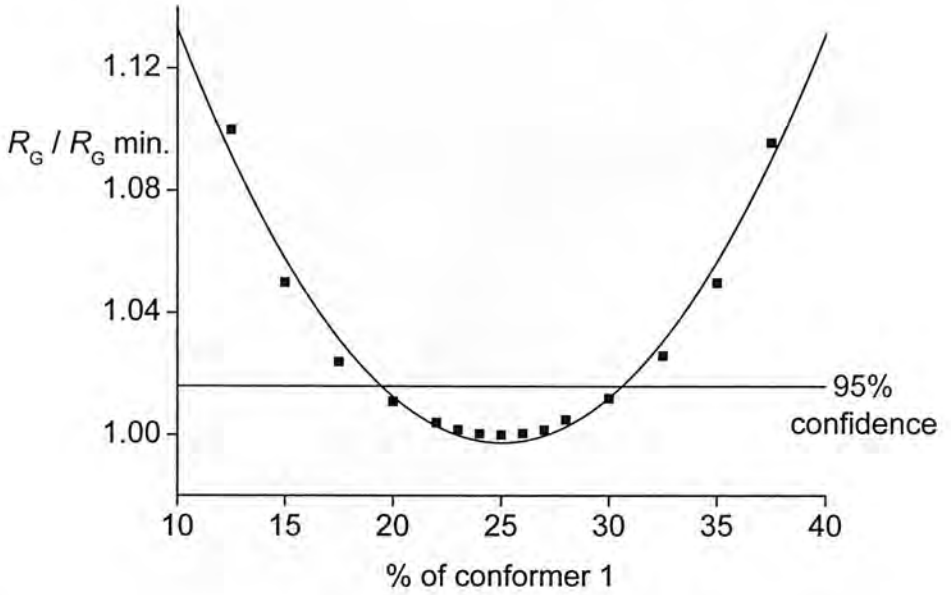


Figure 6.9. R_G / R_G min. for varying amounts of conformer 1 in the SEMTEX refinement of $C(SiMe_2F)_4$. Conformer 2 amount was fixed at 42% for each refinement, with the amount of conformer 3 defined by the amounts of the other two.

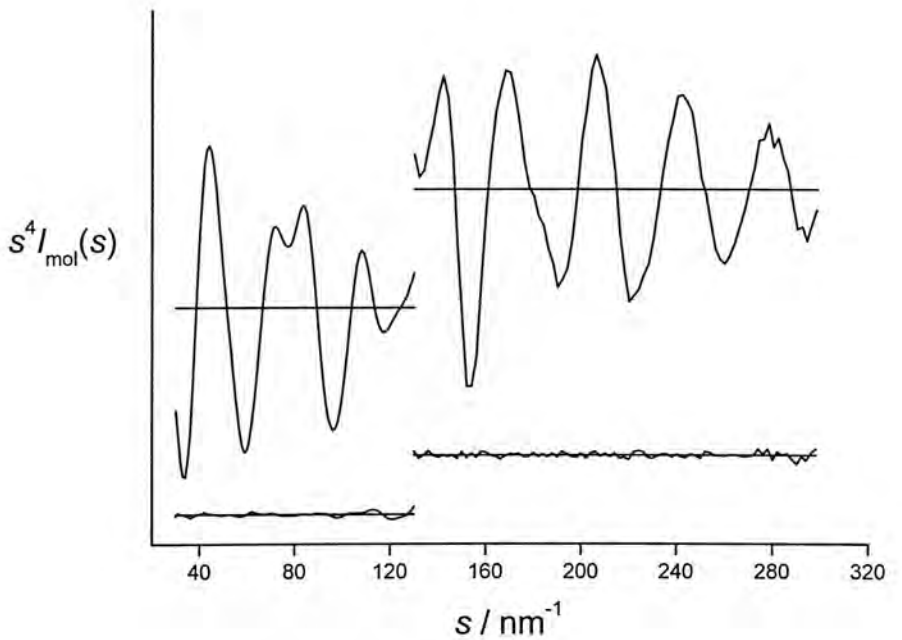


Figure 6.10. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $C(SiMe_2F)_4$.

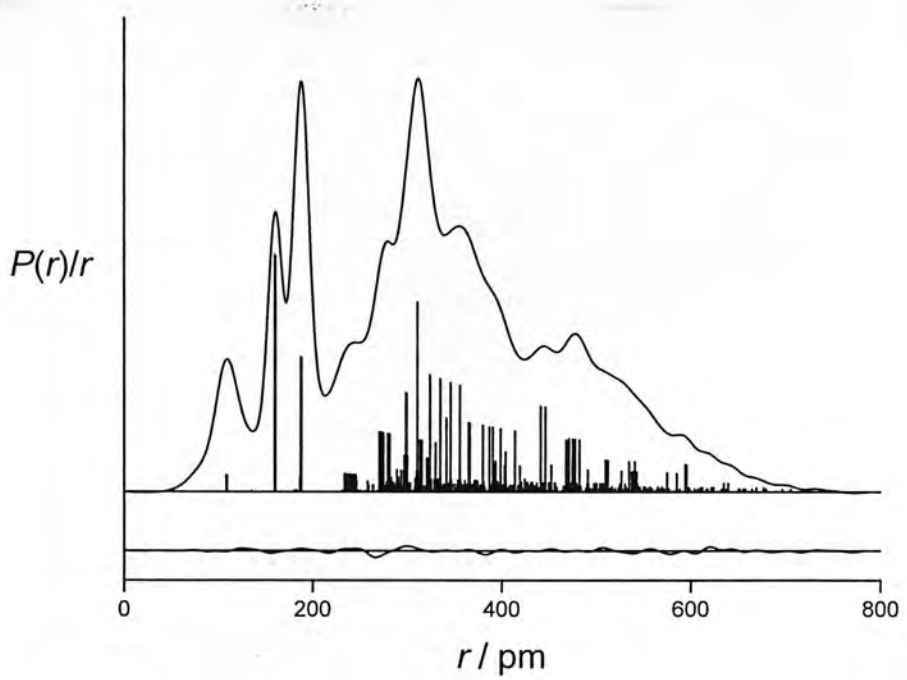


Figure 6.11. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $C(SiMe_2F)_4$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_{Si} - f_{Si})$.

Table 6.11. C-H bond lengths, Si-C-H bond angles and C-Si-C-H bond torsions, Si-F bond lengths and F-Si-C bond angles for $C(SiMe_2F)_4$ calculated with the SEMTEX method. All bond lengths are in pm and angles in °. Interatomic distances are the refined (r_{h1}) values.

Parameter	*r*C-H	∠Si-C-H	∢C-Si-C-H
Conformer 1			
H(7)-C(6)-Si(5)-C(1)	109.2	109.4	59.0
H(8)-C(6)-Si(5)-C(1)	109.5	105.0	178.8
H(9)-C(6)-Si(5)-C(1)	109.3	109.2	-62.9
H(11)-C(10)-Si(5)-C(1)	109.5	105.3	-176.0
H(12)-C(10)-Si(5)-C(1)	109.0	112.8	-53.8
H(13)-C(10)-Si(5)-C(1)	109.5	106.3	66.2
Conformer 2			
H(7)-C(6)-Si(5)-C(1)	109.5	103.8	172.0
H(8)-C(6)-Si(5)-C(1)	109.1	109.7	-68.3
H(9)-C(6)-Si(5)-C(1)	109.2	108.9	53.2
H(11)-C(10)-Si(5)-C(1)	109.5	104.8	-168.7
H(12)-C(10)-Si(5)-C(1)	109.2	111.2	-49.9
H(13)-C(10)-Si(5)-C(1)	109.1	107.1	69.8
Conformer 3			
H(7)-C(6)-Si(5)-C(1)	109.2	109.2	63.7
H(8)-C(6)-Si(5)-C(1)	109.5	104.5	-176.1
H(9)-C(6)-Si(5)-C(1)	109.4	109.0	-57.6

Parameter	*r*C-H	∠Si-C-H	ϕC-Si-C-H
H(11)-C(10)-Si(5)-C(1)	109.5	104.4	-174.6
H(12)-C(10)-Si(5)-C(1)	109.1	108.8	-54.4
H(13)-C(10)-Si(5)-C(1)	109.4	109.2	66.9
Parameter	*r*Si-F	∠C-Si-F	
Conformer 1			
F(39)-Si(4)-C(1)	159.9	107.8	
F(41)-Si(5)-C(1)	160.5	106.5	
Conformer 2			
F(38)-Si(2)-C(1)	159.7	108.4	
F(39)-Si(4)-C(1)	159.9	108.5	
F(40)-Si(3)-C(1)	159.9	107.6	
F(41)-Si(5)-C(1)	160.5	105.7	
Conformer 3			
F(38)-Si(2)-C(1)	160.2	107.3	
F(39)-Si(4)-C(1)	159.8	108.4	
F(40)-Si(3)-C(1)	159.8	106.9	
F(41)-Si(5)-C(1)	159.6	107.2	

Table 6.12. Energy differences and relative abundances of the three conformers, predicted from the HF/6-31G* level of theory and from the SEMTEX GED refinement. The uncertainties for the energy differences are estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 2.

Conformer	Symmetry	Energy difference / kJ mol^{-1}		Abundance / %	
		HF/ 6-31G*	SEMTEX	HF/ 6-31G*	SEMTEX
1	C_2	+ 0.15	0.00	23.0	25.0(31)
2	C_1	0.00	+ 0.59(63)	48.3	42.0(22)
3	C_1	+ 1.78	+ 1.41(96)	28.7	33.0(31)

6.3.3. Tetrakis(bromodimethylsilyl)methane

6.3.3.1. Theoretical methods

The structure of $C(SiMe_2Br)_4$ was determined *ab initio*. A potential-energy surface search involving rotation of the *tert*-butyl groups around the Si-C bonds found three potential conformers which could be present in significant quantity in the experimental mixture: conformer 1, exhibiting C_2 symmetry, and conformers 2 and 3, each exhibiting C_1 symmetry. The molecular geometry of each conformer of $C(SiMe_2Br)_4$ at the MP2/6-311G* level with a LanL2DZ pseudopotential on the bromine atoms can be found in Table 6.3.

The Gibbs free energies of the conformers were calculated at the Hartree-Fock level of theory with the 6-31G* basis set. The energy differences between the three conformers calculated at this level are given in Table 6.13, along with the relative abundances calculated using the Boltzmann distribution.

Table 6.13. Total energies and energy differences between conformers calculated at the HF/6-31G* level. Energy differences are relative to conformer 2, the lowest-energy conformer. Those conformers marked with a * were not included in the refinement model.

Conformer	Symmetry	Total energy / kJ mol^{-1}	Energy difference / kJ mol^{-1}	Abundance / %
1	C_2	-30954931.03	0.00	70.4
2	C_1	-30954921.65	+ 9.4	13.2
3	C_1	-30954921.50	+ 9.5	12.7
4*	D_2	-30954918.20	+12.8	1.4
5*	C_2	-30954906.40	+24.6	0.1
6*	C_1	-30954911.63	+19.8	1.0
7*	C_1	-30954911.62	+19.4	1.0

From these energy differences, therefore, it is possible that all three of these conformers would be present in the experimental mixture, and so three conformers were included in the model written for the least-squares refinement of the structures based on the GED data. Again, the relative abundances were re-calculated based on a three-conformer mixture, giving a predicted ratio of 73.1 : 13.7 : 13.2.

At the outset of the SEMTEX refinement process, the heavy-atom positions were fixed as calculated using SARACEN, and both MP2/LanL2DZ/6-311G* and MM3 calculations were performed on the peripheral atoms. Table 6.14 shows the resulting peripheral-atom parameters from both the MM and MP2 calculations for one Me_2Cl group of each conformer. The full set of peripheral-atom parameters is given in Table S6c.1 in the Supplementary Information..

Table 6.14. Peripheral-atom bond lengths, bond angles and bond torsions for $C(SiMe_2Br)_4$ calculated with the MP2/6-311G* and MM3 methods. A LanL2DZ pseudopotential was used on the bromine atoms in the MP2 calculations. All bond lengths are in pm and angles in °. Internuclear distances are the calculated (r_e) values. In the MM3 calculations, Si atoms were replaced by C atoms. However, only the *differences* between the calculated parameters are relevant to the GED refinements.

Parameter	*r*C-H		∠Si-C-H		∢C-Si-C-H	
Conformer 1	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	109.1	*110.7*	113.9	*112.8*	54.6	*48.9*
H(8)-C(6)-Si(5)-C(1)	109.5	*111.2*	106.8	*109.3*	174.0	*168.0*
H(9)-C(6)-Si(5)-C(1)	109.4	*111.0*	112.2	*110.8*	-67.8	*-74.8*
H(11)-C(10)-Si(5)-C(1)	109.0	*110.8*	113.8	*112.2*	-66.8	*-41.1*
H(12)-C(10)-Si(5)-C(1)	109.4	*110.8*	112.9	*111.6*	56.5	*81.3*
H(13)-C(10)-Si(5)-C(1)	109.6	*111.2*	106.7	*109.8*	175.0	*-160.5*
Conformer 2	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	108.8	*110.2*	114.1	*115.1*	55.9	*58.2*
H(8)-C(6)-Si(5)-C(1)	109.6	*111.2*	105.9	*109.2*	174.2	*178.0*
H(9)-C(6)-Si(5)-C(1)	109.2	*110.9*	111.9	*111.1*	-69.1	*-65.1*
H(11)-C(10)-Si(5)-C(1)	109.7	*111.2*	105.4	*109.7*	-179.5	*-168.5*
H(12)-C(10)-Si(5)-C(1)	108.8	*110.6*	114.8	*113.2*	-62.1	*-49.1*
H(13)-C(10)-Si(5)-C(1)	109.2	*110.6*	113.0	*112.3*	62.9	*73.5*
Conformer 3	MP2	MM3	MP2	MM3	MP2	MM3
H(7)-C(6)-Si(5)-C(1)	109.1	*110.9*	113.4	*111.9*	63.0	*62.6*
H(8)-C(6)-Si(5)-C(1)	109.6	*111.2*	105.3	*109.0*	179.8	*-179.8*

Parameter	*r*C-H		∠Si-C-H		ϕC-Si-C-H	
H(9)-C(6)-Si(5)-C(1)	109.1	*110.6*	113.0	*112.8*	-62.7	*-61.6*
H(11)-C(10)-Si(5)-C(1)	109.5	*111.2*	106.2	*109.8*	-174.1	*-171.0*
H(12)-C(10)-Si(5)-C(1)	109.0	*110.7*	112.9	*112.5*	-56.2	*-51.3*
H(13)-C(10)-Si(5)-C(1)	109.4	*110.9*	112.9	*111.2*	68.3	*70.6*
Parameter						
Conformer 1	MP2	MM3	MP2	MM3		
Br(39)-Si(4)-C(1)	231.3	*226.4*	110.9	*108.6*		
Br(41)-Si(5)-C(1)	232.1	*226.5*	111.8	*108.9*		
Conformer 2	MP2	MM3	MP2	MM3		
Br(38)-Si(2)-C(1)	230.9	*226.1*	112.0	*109.1*		
Br(39)-Si(4)-C(1)	230.7	*226.4*	112.2	*109.3*		
Br(40)-Si(3)-C(1)	231.8	*226.3*	112.3	*109.1*		
Br(41)-Si(5)-C(1)	232.3	*226.1*	110.9	*109.5*		
Conformer 3	MP2	MM3	MP2	MM3		
Br(38)-Si(2)-C(1)	231.6	*226.5*	110.3	*109.1*		
Br(39)-Si(4)-C(1)	230.6	*226.3*	111.6	*109.3*		
Br(40)-Si(3)-C(1)	230.8	*226.4*	110.3	*109.1*		
Br(41)-Si(5)-C(1)	231.4	*226.1*	110.5	*108.8*		

6.3.3.2. SARACEN refinement

The starting parameters for the r_{h1} refinement were taken from the theoretical geometry optimised at the MP2/6-311G* level. Forty-three geometric parameters were refined along with nine groups of vibrational amplitudes. Two parameters to describe the respective amounts of the three conformers in the experimental mixture were also included. Thirty-two geometric and four amplitude restraints were applied according to the SARACEN method, with most of the geometric restraints applying to the torsional parameters for the two less-abundant C_1 conformers. These can be found in Table 6.3. The final R factors for the refinement were found to be R_G = 0.081 and R_D = 0.087. Interatomic distances and corresponding amplitudes of vibration are given in Table S6c.2 in the supplementary information, and final experimental coordinates from the SARACEN GED analysis are given in Table S6c.3. The least-squares correlation matrix is given in Table S6c.4.

6.3.3.3. SEMTEX refinement

The starting parameters were as for the SARACEN refinement. Once all forty-three geometric parameters and nine groups of vibrational amplitudes were refined according to the SARACEN method, the SEMTEX code was activated and the peripheral-atom positions updated computationally using MM. Carbon atoms were again substituted for the silicon atoms for the MM calculations.

As for the SARACEN refinement, all forty-three geometric parameters were refined along with nine groups of vibrational amplitudes. Thirty-two geometric and four amplitude restraints were applied using the SARACEN method. After refinement, the conformational composition was determined. First, the fraction parameter describing the amount of conformer 1 in the sample was altered in order to find which value gave the lowest R factor, with conformers 2 and 3 fixed at equal amounts. The resulting plot of R_G vs $R_{Gmin.}$ is shown in Figure 6.12. As can be seen, the R factor increases either side of the minimum at 57.6%. The 95% confidence limit shows that the uncertainty in this value (~2 estimated standard deviations) is around ±4%. Once the fractional amount corresponding to the best fit was found for

conformer 1, it was fixed and the amount of conformer 2 was altered to find the minimum R factor again, with the amount of conformer 3 defined by the amounts of the other two. The plot of R_G *v* $R_{Gmin.}$ is shown in Figure 6.13 – this time, the uncertainty can be seen to be around ±5%. From these, it can be seen that the best fit comes when the three conformers are present in a 57:22:21 ratio.

In the final refinement, the R factors were $R_G = 0.078$ and $R_D = 0.075$, obtained for a conformer ratio of 57.0 : 22.0 : 21.0 (conf 1 : conf 2 : conf 3). Figure 6.14 shows the molecular scattering intensities for the SEMTEX refinement, while Figure 6.15 shows the radial distribution curve. Table 6.3 lists the final refined parameters, and Table 6.15 shows the resulting peripheral-atom parameters for one Me_2Cl group of each conformer. The relative energies of the conformers calculated from this composition are given in Table 6.16. Interatomic distances and corresponding amplitudes of vibration are given in Table S6c.5 and final experimental coordinates from the SEMTEX GED analysis are given in Table S6c.6. The least-squares correlation matrix is given in Table S6c.7. The full set of peripheral-atom parameters is given in Table S6c.8.

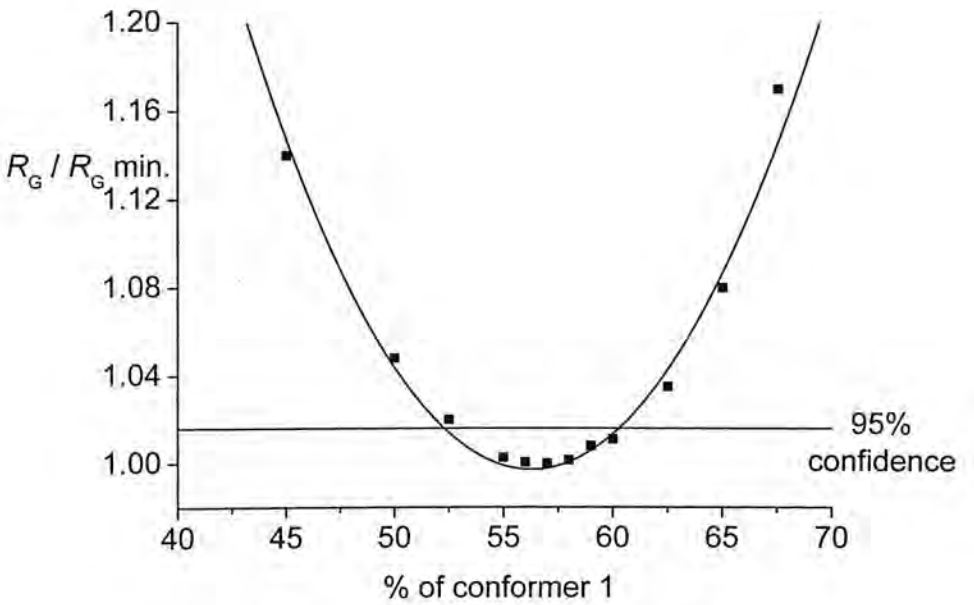


Figure 6.12. R_G / R_G min. for varying amounts of conformer 1 in the SEMTEX refinement of $C(SiMe_2Br)_4$. Conformers 2 and 3 were fixed at equal amounts for each different amount of conformer 1.

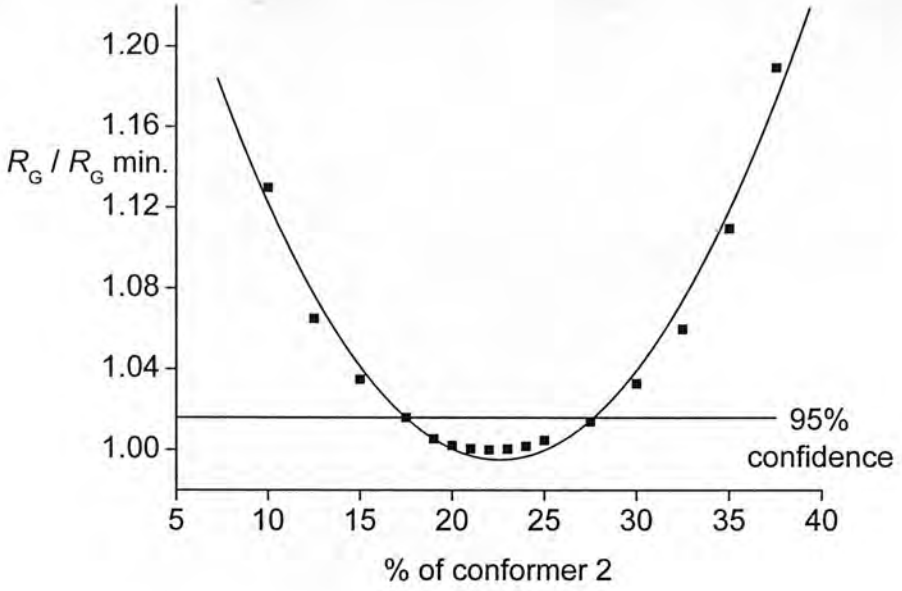


Figure 6.13. R_G / R_G min. for varying amounts of conformer 2 in the SEMTEX refinement of $C(SiMe_2Br)_4$. Conformer 1 amount was fixed at 75% for each refinement, with the amount of conformer 3 defined by the amounts of the other two.

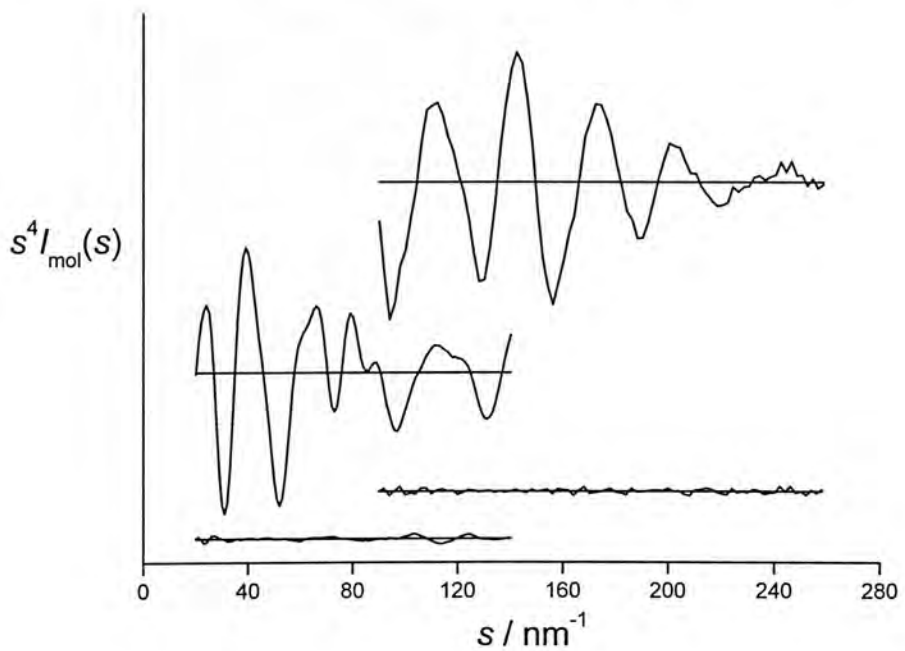


Figure 6.14. Experimental and weighted difference (experimental – theoretical) molecular scattering intensities for $C(SiMe_2Br)_4$.

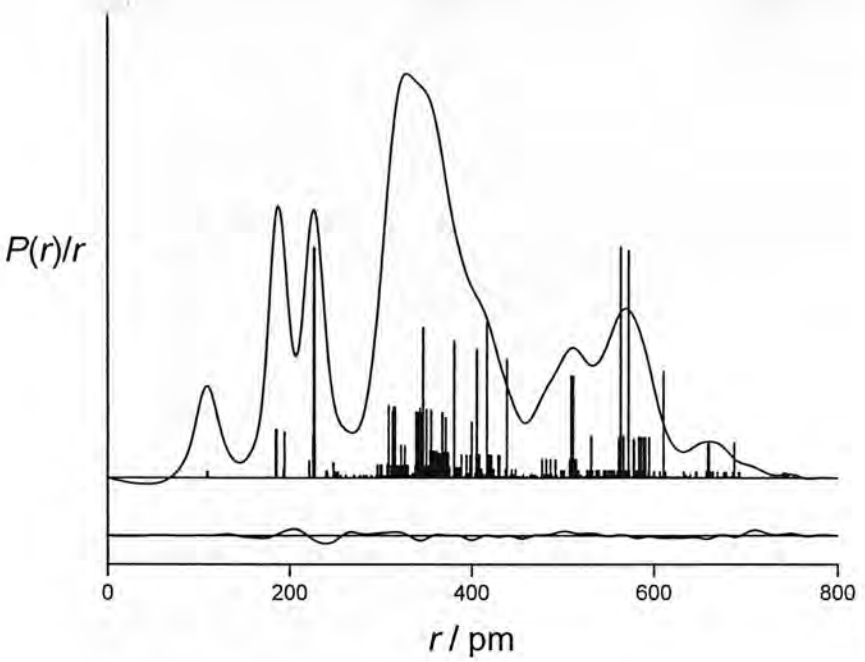


Figure 6.15. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SEMTEX refinement of $C(SiMe_2Br)_4$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_C - f_C)(Z_{Br} - f_{Br})$.

Table 6.15. C-H bond lengths, Si-C-H bond angles and C-Si-C-H bond torsions, Si-Br bond lengths and Br-Si-C bond angles for $C(SiMe_2Br)_4$ calculated with the SEMTEX method. All bond lengths are in pm and angles in °. Interatomic distances are the refined (r_{h1}) values.

Parameter	rC-H	∠Si-C-H	ϕC-Si-C-H
Conformer 1			
H(7)-C(6)-Si(5)-C(1)	109.5	114.9	59.5
H(8)-C(6)-Si(5)-C(1)	110.0	107.7	178.6
H(9)-C(6)-Si(5)-C(1)	109.8	113.2	-64.2
H(11)-C(10)-Si(5)-C(1)	109.4	114.8	-59.8
H(12)-C(10)-Si(5)-C(1)	109.9	113.9	62.6
H(13)-C(10)-Si(5)-C(1)	110.0	107.6	-179.3

Parameter	rC-H	∠Si-C-H	ϕC-Si-C-H
Conformer 2			
H(7)-C(6)-Si(5)-C(1)	109.2	115.1	59.4
H(8)-C(6)-Si(5)-C(1)	110.0	106.8	179.1
H(9)-C(6)-Si(5)-C(1)	109.6	112.9	-64.0
H(11)-C(10)-Si(5)-C(1)	110.1	106.3	176.8
H(12)-C(10)-Si(5)-C(1)	109.2	115.8	-63.7
H(13)-C(10)-Si(5)-C(1)	109.6	114.0	58.9
Conformer 3			
H(7)-C(6)-Si(5)-C(1)	109.5	114.5	60.7
H(8)-C(6)-Si(5)-C(1)	110.1	106.2	178.3
H(9)-C(6)-Si(5)-C(1)	109.5	114.1	-63.5
H(11)-C(10)-Si(5)-C(1)	110.0	107.1	-177.9
H(12)-C(10)-Si(5)-C(1)	109.4	113.9	-58.2
H(13)-C(10)-Si(5)-C(1)	109.8	113.9	63.7
Parameter			
Conformer 1			
Br(39)-Si(4)-C(1)	226.2	109.1	
Br(41)-Si(5)-C(1)	227.0	110.0	
Conformer 2			
Br(38)-Si(2)-C(1)	225.9	110.1	

Parameter	rC-H	∠Si-C-H	ϕC-Si-C-H
Br(39)-Si(4)-C(1)	225.6	110.3	
Br(40)-Si(3)-C(1)	226.8	110.4	
Br(41)-Si(5)-C(1)	227.3	109.0	
Conformer 3			
Br(38)-Si(2)-C(1)	226.5	108.5	
Br(39)-Si(4)-C(1)	225.6	109.8	
Br(40)-Si(3)-C(1)	225.8	108.5	
Br(41)-Si(5)-C(1)	226.4	108.7	

Table 6.16. Energy differences and relative abundances of the three conformers, predicted from the HF/6-31G* level of theory and from the SEMTEX GED refinement. The uncertainties for the energy differences are estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 1.

Conformer	Symmetry	Energy difference / kJ mol^{-1}		Abundance / %	
		HF/ 6-31G*	SEMTEX	HF/ 6-31G*	SEMTEX
1	C_2	0.00	0.00	73.1	57.0(20)
2	C_1	+ 9.38	+ 6.52(65)	13.7	22.0(28)
3	C_1	+ 9.53	+ 6.70(67)	13.2	21.0(28)

6.4. Discussion

The molecular structures of tetrakis(-chlorodimethylsilyl)methane, tetrakis(-fluorodimethylsilyl)methane and tetrakis(bromodimethylsilyl)methane have been determined in the gas phase using both the SARACEN and SEMTEX methods. For the SEMTEX method, the molecular mechanics calculations (the 'DYNAMITE' part) were performed using carbon atoms substituted for the silicon atoms, since none of the force fields associated with the TINKER package could handle Si-X bonds. At every stage of the refinement, however, the peripheral-atom parameters returned by MM are immediately modified by a set of differences between MM and MP2 calculations, done with the core atoms frozen to GED positions. This substitution therefore has no effect on the final refined structure of the molecule. A comparison of the peripheral-atom parameters from the MP2 calculations and from the final SEMTEX refinements shows that in the final refinement, the differences in peripheral-atom parameters are being correctly applied from the MP2 calculations, with the average values still determined by GED.

Previous ^{1}H and ^{29}Si NMR study of $C(SiMe_2Cl)_4$ and $C(SiMe_2Br)_4$ in solution had found that at low temperature the compounds are present as a mixture of C_1 and C_2 conformers, with a much greater abundance of C_1 (C_1:C_2 ~ 85:15 for both cases). The computed result that other conformers of higher symmetry would be too energetically unfavourable to exist in quantities significant to experimental study is therefore in agreement with this observation. For each molecule, a three-conformer model gave a much better fit to the electron-diffraction data than any of the three single-conformer models, vindicating the computed result that all three conformers were present in significant quantities in the experimental sample. From the refinements, it was found that in the gas-phase sample, the $C(SiMe_2Br)_4$ mixture was composed of 57% C_2 conformer, a considerable difference between the gas-phase structure and that found by NMR in solution. It was indicated from the NMR spectra that at higher temperatures, it would be possible for internal rotation of the $SiMe_2X$ groups to lead to interconversion between different conformers of C_1 symmetry. Indeed, the potential-energy scans conducted at the outset of the present work suggested the possibility of four different C_1 conformers. However, it was found that

two of the possible C_1 conformers would be very energetically unfavourable, and so only two were used in the GED refinement model. Interconversion between these two conformers would be most difficult in the case of $C(SiMe_2Br)_4$, where the large bromine atoms act as a hindrance to internal rotation, while $C(SiMe_2Cl)_4$ would be more likely to exhibit this behaviour, as would $C(SiMe_2F)_4$ which not was included in the NMR study,

In the refinement of $C(SiMe_2Cl)_4$, the presence of substantially more conformer 1 (with C_2 symmetry) than either of the two conformers with C_1 symmetry illustrates that this conformer was considerably lower in energy, as predicted from the *ab initio* calculations. The calculated relative energies of the three conformers, given in Table 6.5., indicated a mixture comprising 70.5% of conformer 1, 15.5% of conformer 2 and 14% conformer 3. This is therefore different from the final conformational composition obtained from the SEMTEX refinement, of 75.0(20)% conformer 1, 8.0(28)% conformer 2 and 17.0(28)% conformer 3, corresponding to the relative energies given in Table 6.8. As can be seen by comparison of the values in Tables 6.4 and 6.8, the largest difference is in the relative abundances of conformers 2 and 3. The HF/6-31G* calculations predict that conformer 2 should be slightly the more abundant of the two, whereas the GED refinement gives conformer 3 more than twice as abundant as conformer 2. The uncertainty in the energy differences between conformers were estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 1 (see Figure 6.4).

For both the SARACEN and SEMTEX refinements of $C(SiMe_2Cl)_4$, the core-atom bond length parameters are different between the experimental and calculated structures. The average bond length rC-Si av. refines to 190.8 pm with SARACEN and 191.0 pm using SEMTEX, a differences of 1.8 pm and 2.0 pm respectively from the value of 189.0 pm calculated using MP2/6-311G*. However, it is common for first- and second-row to be calculated ~ 2 pm out by this level of theory, and so this discrepancy is readily explained. More significantly, the difference parameter, rC-Si diff., refines to values of -0.6 pm and -0.4 pm respectively, compared with the MP2 value of 0.5 pm – this indicates that the $SiMe_2Cl$ group-carbons have longer C-Si bonds than the central C atom, rather than shorter as predicted by MP2.

For the core-atom angles, varying degrees of agreement between theory and experiment are observed for both the SARACEN and SEMTEX refinements. The average Si-C-Si angle refines to 109.8° using SEMTEX and 111.8° using SARACEN, compared with a value of 109.4° from *ab initio*, an example where SEMTEX matches the calculated structure much more closely than SARACEN. In contrast, the average C-Si-Me angle is given as 117.6° by SARACEN and 113.3° by SEMTEX, with the calculated value of 115.0° lying in between. It should, however, be noted that in this case too, the SEMTEX value is considerably closer to that of the *ab initio* structure than the SARACEN value – 1.7° as opposed to 2.6°.

The $SiMe_2Cl$-group torsion parameters could be refined without restraints only for conformer 1, the most abundant conformer. For this case, ($SiMe_2Cl$) tor1 (p_{16}) refined to 70.8° and 83.3° using SARACEN and SEMTEX respectively, a difference of 12.5°. When these are compared with the MP2 value of 75.0°, it can be seen that the group rotates in a different direction relative to the calculated structure depending on which refinement method is used. For $SiMe_2Cl$ tor2 (p_{17}), in contrast, the values from both refinements were very similar, 18.2° from SARACEN and 18.0° from SEMTEX. In this case, both were very different to the MP2 value of -40.3°.

Both methods give a very similar value for the average Si-Cl distance: 208.6 pm by SARACEN and 208.5 pm by SEMTEX, while the MP2 value is again lower, at 207.0 pm. In the case of SARACEN, however, this represents the value for every Si-Cl bond in the refined structure. In contrast, using SEMTEX, each of the (non-symmetry-equivalent) chlorine atoms is able to have a different Si-Cl distance, the differences between these being determined *ab initio*. Thus, the Si-Cl distances take a range of 1.3 pm, from 107.8 pm to 109.1 pm, as seen in Table 6.7. The inclusion of *ab initio* differences has even more marked consequences for the C-Si-Cl angles, which are all constrained to the refined value of 110.1° for ∠C-Si-Cl *av.* using SARACEN. Using SEMTEX, these take a range of 4.2° around the refined average of 110.8°.

However, the advantage of using SEMTEX over SARACEN is most clearly illustrated by its effect on the hydrogen atom positions, and consequently the structure of the methyl groups. In the case of SARACEN, the C-H distances and Si-C-H angles are constrained to the same values for all 72 H atoms: *r*C-H *av.* and ∠Si-

C-H *av.*, which refine to 107.2 pm and 111.1° respectively. Furthermore, all three H atoms within each methyl group are described using a single methyl torsion. Thus, all methyl groups are constrained to C_{3v} symmetry. For SEMTEX, none of these constraints apply. This is especially important in the case of the Si-C-H angles, which take a very large range of values – 17.0° – about the refined average value of 110.9°.

The SEMTEX method clearly gives a considerably different structure to that obtained by the SARACEN method for this molecule. A comparison of the *R*-factor values obtained using the two methods indicates that the refined structure given by SEMTEX provides a considerably better fit to the experimental data – $R_G = 0.089$ for SEMTEX, $R_G = 0.106$ for SARACEN. This is likely to be due to two factors. Using SEMTEX, the $SiMe_2Cl$ groups within each conformer are able to refine to a completely asymmetric conformation if so desired (as indicated by *ab initio* calculations), increasing the likelihood of a good match between the refined structure and the GED data. Also, the values of average peripheral-atom distances and angles within each separate conformer are no longer constrained to the average as calculated over all three conformers. As can be seen in Table 6.7, for example, the average Si-Cl distance using SEMTEX is 209.0 pm in conformer 1 and 208.4 pm in conformer 3, whereas in SARACEN these would be constrained to take an identical value for all three conformers in the absence of extra refining parameters.

For tetrakis(fluorodimethylsilyl)methane, the conformational composition of 25.0(31) : 42.0(22) : 33.0(31) (conf. 1 : conf. 2 : conf. 3) obtained from the SEMTEX refinement is in reasonable agreement with the composition of 23.0 : 48.3 : 28.7 predicted from the *ab initio* calculations. After taking into account the double degeneracy of the C_2 conformer (conformer 1) relative to the other two conformers, it can be seen that the refinement indicates very small energy differences between the conformers: conformer 3, the least energetically favourable, is only 1.4 kJ mol^{-1} higher in energy than conformer 1. The uncertainty in the energy differences between conformers were estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 1 (see Figure 6.8).

For both the SARACEN and SEMTEX refinements of $C(SiMe_2F)_4$, the core-atom bond length parameters are in varying degrees of agreement with those calculated *ab initio*. The average bond distance *r*C-Si av. refines to 187.2 pm with SARACEN and 187.4 pm using SEMTEX, very similar to the value of 187.5 pm calculated using MP2/6-311G*. In contrast the difference parameter, *r*C-Si diff., refines to values of 0.9 pm and -0.4 pm respectively, compared with the MP2 value of 3.0 pm.

For the core-atom angles, notable differences between theory and experiment are observed for both the SARACEN and SEMTEX refinements. The average Si-C-Si angle refines to 107.5° using SEMTEX and 108.0° using SARACEN, compared with a value of 109.9° from *ab initio*, a substantial difference between theory and experiment. Similarly, the average C-Si-Me angle is given as 116.1° by SARACEN and 116.0° by SEMTEX, with the calculated value of 114.5° considerably smaller.

As for $C(SiMe_2Cl)_4$, the $SiMe_2F$-group torsion parameters could be refined without restraints only for conformer 1, the most abundant conformer. For this case, ($SiMe_2F$) tor1 (p_{16}) refined to values of 48.9° and 63.4° using SARACEN and SEMTEX respectively, a difference of 14.5°. When these are compared with the MP2 value of 43.6°, it can be seen that the SEMTEX method predicts a much greater rotation of this group away from the calculated structure than that predicted by SARACEN. For ($SiMe_2F$) tor2 (p_{17}), in contrast, the values from both refinements were very similar, -82.0° from SARACEN and -80.3° from SEMTEX, both similar to the computed value of -76.8°.

The two methods give differing values for the average Si-F distance: 160.5 pm by SARACEN and 160.0 pm by SEMTEX, while the MP2 value is very significantly higher, at 162.5 pm. For the SARACEN refinement, this represents the value for every Si-F bond in the refined structure. In the case of SEMTEX however, each of the fluorine atoms is able to have a different Si-F distance, as predicted from the MP2 calculations. The Si-F distances take a range of 0.8 pm, from 159.6 pm to 160.5 pm, as seen in Table 5.11. Likewise, the C-Si-F angles are no longer all constrained to the refined value of 105.4° for ∠C-Si-F *av.* found using SARACEN: applying SEMTEX, these take a range of 2.8° around the refined average of 106.2°.

Using SEMTEX in place of SARACEN also has a marked effect on the hydrogen atom positions – the methyl groups are no longer constrained to C_3 symmetry, and

each individual H-atom is described by a unique distance, angle and torsion. For SARACEN, *r*C-H *av.* and ∠Si-C-H *av.* refine to 109.6 pm and 106.2° respectively, and describe the positions of all 72 H atoms in the structure. For SEMTEX, the C-H distances cover a modest range of 0.6 pm, around the refined *r*C-H *av.* value of 109.3 pm. The Si-C-H angles, in contrast, take a much larger range of values – 9.2 ° – about the refined average value of 109.8°.

Comparing the *R*-factor values obtained using the two methods indicates that the refined structure given by SEMTEX provides a considerably better fit to the experimental data – R_G = 0.059 for SEMTEX, R_G = 0.071 for SARACEN. As for the previous $C(SiMe_2Cl)_4$ refinement, this is likely to be a consequence of the fact that each methyl group within a $SiMe_2F$ group is allowed to refine to a completely asymmetric conformation in the SEMTEX refinement, as indicated by *ab initio* calculations. The values of average peripheral-atom distances and angles can also vary between the different conformers using SEMTEX, no longer constrained to an identical value for all three conformers as they are in the SARACEN refinement.

For tetrakis(bromodimethylsilyl)methane, the conformational composition of 57.0(20) : 22.0(28) : 21.0(28) (conf. 1 : conf. 2 : conf. 3) obtained from the SEMTEX refinement is notably different from the composition of 73.1 : 13.7 : 13.2 predicted from the *ab initio* calculations. Again accounting for the double degeneracy of the C_2 conformer (conformer 1) relative to the other two conformers, the refinement indicates that conformer 1 is lowest in energy, with conformers 2 and 3 predicted at 6.38 kJ mol^{-1} and 6.53 kJ mol^{-1} higher in energy respectively. In comparison with the values calculated *ab initio*, it can be seen that the refinement predicts conformers 2 and 3 to be considerably closer in energy to conformer 1. The uncertainties in the energy differences between conformers were estimated using the Boltzmann distribution from the energy difference at ±1 standard deviation about the optimum percentage of conformer 1 (see Figure 5.13).

The core-atom bond length parameters are again in varying degrees of agreement with those calculated *ab initio* for both the SARACEN and SEMTEX refinements. The average bond distance *r*C-Si av. refines to 187.5 pm with SARACEN and 187.8 pm using SEMTEX, again very similar to the value of 187.5 pm calculated using MP2/6-311G*. The difference parameter *r*C-Si diff. refines to 7.7 pm and 8.6 pm,

respectively, again considerably different from the MP2 value of 5.0 pm. This time however, the refined values indicate a larger difference in Si-C distances than calculated *ab initio*, rather than a much smaller one as found for $C(SiMe_2F)_4$.

For the core-atom angles, the differences between theory and experiment are less pronounced, but slight differences are noticeable between the SARACEN and SEMTEX methods. The average Si-C-Si angle refines to 103.5° using SARACEN, identical to the computed value. Using SEMTEX, a slight difference is observed, as the parameter refines to 105.0°. Conversely, the average C-Si-Me angle is given as 114.8° by SEMTEX, very close to the computed value of 114.9°, while this time the SARACEN refinement predicts a higher value of 116.4°.

For the conformer 1 $SiMe_2Br$-group torsion parameters, $(SiMe_2Br)tor1$ (p_{16}) refined to values of 48.5° and 54.1° using SARACEN and SEMTEX respectively, a small difference of 5.6°. When these are compared with the MP2 value of 39.4°, it can be seen that much less deviation is found between theory and experiment here than for the analogous parameter in the $C(SiMe_2F)_4$ refinement. For $SiMe_2Brtor2$ (p_{17}) the values from both refinements were again very similar, -80.3° from SARACEN and -78.0° from SEMTEX, both also close to the computed value of -74.5°. Again, the other $SiMe_2Br$-group torsion parameters could not be refined without applying restraints.

Similar values are obtained from the two methods for the average Si-Br distance: 226.5 pm by SARACEN and 226.3 pm by SEMTEX, while the MP2 value in this case is 1.5 pm lower, at 225.0 pm. In the SEMTEX refinement, the Si-Br distances take a range of 1.7 pm, from 225.6 pm to 227.3 pm, as seen in Table 6.16 – a much larger range than was found for the Si-F distances in $C(SiMe_2F)_4$. In contrast, the C-Si-Br angles show considerably less variation than the C-Si-F angles in $C(SiMe_2F)_4$: a range of 1.9° around the refined average of 108.8°, while the average value for the SARACEN refinement refines to 108.7°, almost identical to that from SEMTEX.

Using SEMTEX in place of SARACEN again has a substantial effect on the structure of the methyl groups. In this case, the SARACEN refinement gives *r*C-H *av.* and aSi-C-H *av.* values of 109.5 pm and 111.6° respectively. For SEMTEX, the C-H distances cover a range of 0.9 pm, around the refined *r*C-H *av.* value of 109.3 pm.

The Si-C-H angles again take a large range of values – 9.9 ° – about the refined average value of 111.9°.

For the refinement of $C(SiMe_2Br)_4$, *R*-factor values obtained using the two methods are $R_G = 0.078$ for SEMTEX, and $R_G = 0.081$ for SARACEN. It can therefore be seen that the application of the SEMTEX method makes less of an improvement to the overall goodness-of-fit than for the $C(SiMe_2F)_4$ and $C(SiMe_2Cl)_4$ cases. One reason for this could be the possible increased influence of the phase effect on the electron diffraction data obtained for $C(SiMe_2Br)_4$ – as bromine is much heavier than either chlorine or fluorine, this could cause a significant distortion of the data in this case. The use of a pseudopotential on the bromine atoms (which was not necessary for the chlorinated or fluorinated analogues) may also have had an effect on the accuracy of the *ab initio* bromine-atom parameters used in the SEMTEX refinement.

6.5. Conclusion

The structures of the molecules $C(SiMe_2Cl)_4$, $C(SiMe_2F)_4$ and $C(SiMe_2Br)_4$ have been examined in the gas phase using both the SARACEN and SEMTEX methods. In each case, a three-conformer model was found to give the best fit to the data, with the C_2 conformer most abundant for $C(SiMe_2Cl)_4$ and $C(SiMe_2Br)_4$, and a C_1 conformer most abundant for $C(SiMe_2F)_4$. The SEMTEX method has thus been applied for the first time to a structure with more than one type of peripheral atom – in this case, H and X [X = Cl, F, Br]. In each case, implementation of the SEMTEX method led to a considerably different structure than that obtained using SARACEN. The use of SEMTEX led to a significant improvement in the quality of fit to the GED data for both $C(SiMe_2Cl)_4$ and $C(SiMe_2F)_4$, but made less of a difference to the overall goodness-of-fit for the structure of $C(SiMe_2Br)_4$, possibly a consequence of the use of a pseudopotential on the bromine atoms in the *ab initio* calculations. The results of these refinements show that the method is of considerable use in the refinement of such challenging, multi-conformer structures, and can be applied with confidence to molecules with several different types of peripheral atom.

6. 6. References

1. C. Eaborn and P. D. Lickiss, *J. Organomet. Chem.*, 1985, **294**, 305.
2. B. Beagley, R. G. Pritchard, and J. O. Titiloye, *J. Mol. Struct.* 1989, **212**, 323; B. Beagley, R. G. Pritchard and J. O. Titiloye, *J. Mol. Struct.* 1988, **176**, 81.
3. D. Iroff and K. Mislow, *J. Amer. Chem. Soc.*, 1978, **100**, 2121.
4. Buerger, U. Goetze, W. Sawodny, *Spectrochim. Acta, Part A: Molecular and Biomolecular Spectroscopy*, 1970, **26**, 685.
5. K. Bätz, S. L. Masters and P. D. Lickiss, Personal correspondence.
6. D. G. Anderson, D. W. H. Rankin, H. E. Robertson, A. H. Cowley, and M. Pakulski, *J. Mol. Struct.* 1989, **196**, 21.
7. C. A. Morrison, D. W. H. Rankin, H. E. Robertson, P. D. Lickiss and P. C. Masangane, *J. Chem. Soc., Dalton Trans.*, 1999, 2293.
8. P. J. Hay and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 270.
9. W. R. Wadt and P. J. Hay, *J. Chem. Phys.* 1985, **82**, 284.
10. P. J. Hay, and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 299.
11. W. C. Hamilton, *Acta Crystallogr.* 1965, **18**, 502.

7. The gas-phase structure of $Fe_3(CO)_{12}$

7.1. Introduction

The tri-nuclear metal cluster molecule triirondodecacarbonyl [$Fe_3(CO)_{12}$] has a multitude of applications in catalytic processes, ranging from the derivation of catalysts for the hydrogenation of ketones[1] to the process of heavy crude-oil upgrading.[2] The molecular structure of $Fe_3(CO)_{12}$ has been well characterised in the crystalline state[3], but considerably less is known about its structure in the gas phase. Previous theoretical studies[4-6] on the family of $M_3(CO)_{12}$ molecules (M = Fe, Ru, Os) have shown $Fe_3(CO)_{12}$ to adopt a different isomeric structure when compared with the analogous heavier-metal compounds $Ru_3(CO)_{12}$ and $Os_3(CO)_{12}$ – a single structure with C_{2v} symmetry, as opposed to a D_{3h}-symmetry structure as found for the two heavier-metal clusters. The molecule adopts such a structure as the result of the formation of two Fe-C(O)-Fe bridges between one pair of iron atoms.[5] This results in several different Fe-C and C=O distances within the structure, making such a bridging system an interesting and challenging candidate for GED structure determination.

It was initially hoped that the DYNAMITE and SEMTEX methods could also be used in this investigation, applied to the peripheral oxygen atoms. However, time constraints prevented this from being possible: it was determined that an extensive re-write of the SEMTEX code would be necessary in order to deal with the high co-ordination numbers of the metal centres, since the TINKER molecular mechanics code used in the method only provides for atoms with coordination numbers of up to four.

7.2. Experimental

7.2.1. Computational methods

A search of the potential-energy surface revealed two possible structures of $Fe_3(CO)_{12}$ as expected, one with C_{2v} symmetry, containing ten terminal carbonyl ligands and two Fe-(CO)-Fe bridges, and the other with D_{3h} symmetry and no bridging carbonyl groups. However, previous work has shown that the energy difference between these structures in the gas phase is so great that only the C_{2v} form

is expected to be present in any significant quantity in the experimental mixture: Hunstock *et al*[4] report an energy difference of ~ 28.0 kJ mol^{-1} from DFT calculations, which indicates the well over 99.9% of the experimental mixture would be C_{2v} conformer. Geometry optimisations were therefore carried out on the C_{2v} conformer only. A LANL2DZ pseudopotential[7-9] was used on the Fe atoms.

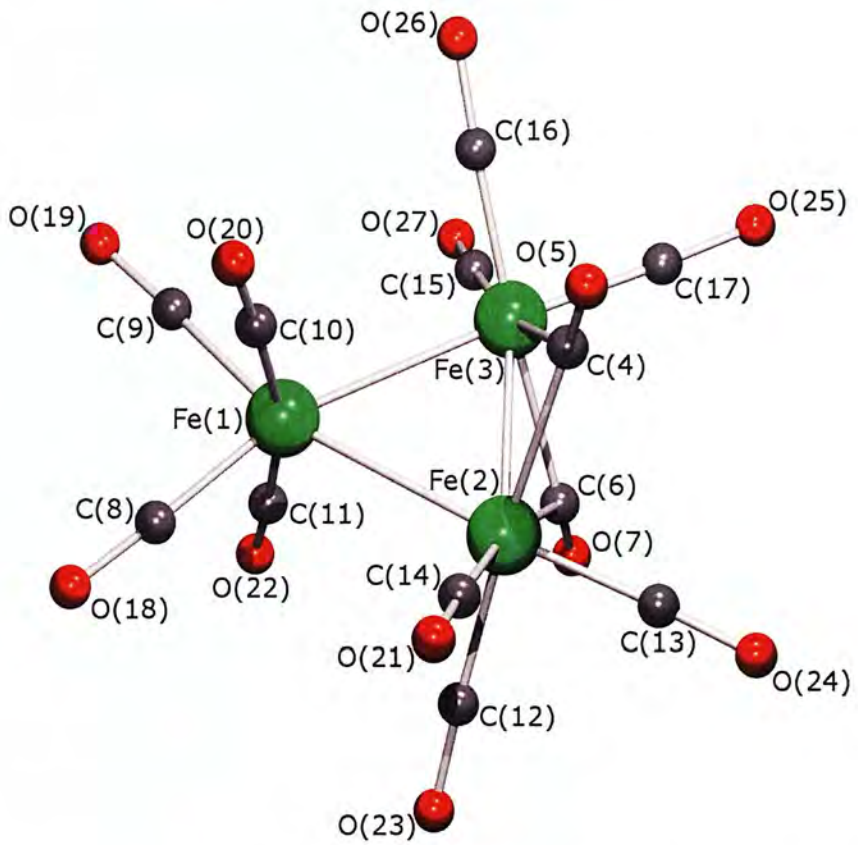


Figure 7.1. Gas-phase molecular structure of $Fe_3(CO)_{12}$. The structure exhibits C_{2v} symmetry.

Amplitudes of vibration (u_{h1}) and curvilinear vibrational correction factors (k_{h1}) to distances required for the GED refinement were derived as described in Chapter 2.

7.2.2. Gas-phase electron diffraction

The Edinburgh gas-phase electron diffraction (GED) apparatus was used to collect data from a sample of $Fe_3(CO)_{12}$ purchased from Aldrich and used without further

purification. The weighting points for the off-diagonal weight matrices, correlation parameters and scale factors for both distances are given in Table S7.1.

7.2.3. Electron diffraction model

To describe the structure of $Fe_3(CO)_{12}$ a model with C_{2v} symmetry was written using twelve geometric parameters, comprising five bond lengths, five bond angles and two bend parameters to describe the position of the bridging CO groups relative to the core Fe atoms.

The Fe–Fe distances were described by the average and a difference (p_{1_2}), defining the difference between the two longer Fe–Fe distances and the shorter Fe(2)-Fe(3) distance. An average and a difference parameter were also used to describe the Fe–C distances (p_{3_4}), with the difference parameter in this case describing the difference between the typical axial/equatorial Fe-C bond length, and the bond length from Fe(2)/Fe(3) to one of the bridging C atoms. A positive value in this case indicates that the distance to one of the bridging C atoms is longer. *Ab initio* calculations showed that while in principle there are five different Fe–C distances for the non-bridging C atoms, the differences are so small that there would be no benefit in using more than one parameter to describe all five of these. A further two parameters were used to describe C=O bonded distances. These were an overall average *r*C=O av. (p_5), and a difference parameter rC=O diff (p_6), describing the difference between the terminal and bridging C=O groups, with a positive value indicating that the bridging C=O distances are longer.

Six bond angles were also used to describe the structure. The average and difference between the C–Fe(1)–Fe angles were used to describe the positions of the axial and equatorial C=O groups bonded to Fe(1) (p_7 and p_8). To describe the positions of the non-bridging C=O groups on Fe(2) and Fe(3), three angle parameters were used. The angle C-Fe-Fe(1)planar (p_9) was used to describe the positions of the C=O groups lying in the plane of the Fe triangle, and the angles C–Fe(2)–Fe(1) (p_{10}) and C–Fe(2)-Fe(3) (p_{11}) describe the positions of the other carbon atoms, not lying in this plane. An angle parameter O=C–Fe was also used to describe the positions of the oxygen atoms relative to the C–Fe bonds (p_{12}).

Finally, two further parameters were used to describe the positions of the bridging C=O groups. A bridging angle (p_{13}) describing the bend of the C-atoms out of the plane of the Fe-triangle was used, defined as the angle between C-A-X, where A is the midpoint of Fe(2) and Fe(3), and X an arbitrary point in the positive direction along the x-axis. Along with this, an O=C...Fe(1) angle (p_{14}) was used, describing the degree of planarity of the bridging oxygen atoms relative to a plane containing the corresponding bridging carbon and the two Fe atoms, with a larger value indicating a greater bend away from the Fe(1) atom.

7.3 Results

7.3.1 Calculations

The molecular structure of $Fe_3(CO)_{12}$ was determined *ab initio*. The molecular geometry at the B3PW91/6-311+G* and MP2/6-311+G* levels is shown in Table 7.1. The MP2 calculation was found to be very difficult to converge well. MP2 is known to perform poorly when transition metals are present, and this can be seen to be the case here. The MP2 calculation predicts an obtuse bridge angle of 108.9°, and very low values for the Fe-Fe average distance and the Fe-C average distance. Results of previous studies[3] indicate that the geometry obtained from the B3PW91 calculation is much closer to the true geometry, and therefore the starting values used for the refinement were those taken from the B3PW91/6-311+G* level of theory.

Table 7.1. Molecular geometry of $Fe_3(CO)_{12}$ calculated using the B3PW91 and MP2 methods. The 6-311+G* basis set was used on all atoms. All distances are in pm and all angles are in degrees.

Parameter		B3PW91[b] (r_e)	MP2
p_1	*r*Fe–Fe average	262.2	240.8
p_2	*r*Fe–Fe difference	13.8	27.3
p_3	*r*Fe-C average	188.7	178.2
p_4	*r*Fe-C difference	18.6	28.5

Parameter		B3PW91[b] (r_e)	MP2
p_5	rO=C average	115.3	118.0
p_6	rO=C difference	2.3	4.6
p_7	∠C-Fe(1)-Fe average	94.2	96.7
p_8	∠C-Fe(1)-Fe difference	14.2	-53.4
p_9	∠C-Fe(2)-Fe(1) planar	175.1	157.1
p_{10}	∠C-Fe(2)-Fe(1)	89.6	109.9
p_{11}	∠C-Fe(2)-Fe(3)	125.5	137.5
p_{12}	∠O=C-Fe	176.5	172.4
p_{13}	bridge angle	66.0	108.9
p_{14}	∠O(bridge)-C…Fe(1)	138.5	127.6
Dependent parameters			
rFe(1)-Fe(2/3)		269.1	254.5
rFe(2)-Fe(3)		255.3	227.1
rFe-C average (terminal)		179.4	163.9
rFe-C (bridging)		198.0	192.5
rO=C average (terminal)		114.1	115.7
rO=C (bridging)		116.5	120.3

7.3.3. SARACEN refinement

The starting values for the twelve geometric parameters used in the refinement were taken from the highest-level DFT calculation carried out (B3PW91/LanL2DZ/6-

311+G*). The model was refined as an r_{h1} structure (using curvilinear amplitude corrections). In total, all 12 geometric parameters and 10 groups of amplitudes were refined. Where parameters were not sufficiently well defined to refine to sensible values restraints were applied in accordance with the SARACEN method.

The final R factor for the refinement was $R_G = 0.081$ ($R_D = 0.037$). Final refined parameters are listed in Table 7.2. Figure 7.2 shows the molecular-scattering intensity curve from the refinement. The radial-distribution curve for the refinement is shown in Figure 7.3. Interatomic distances and the corresponding amplitudes of vibration are given in Table S7.1, with the final experimental coordinates for the GED analysis given in Table S7.2. The least-squares correlation matrix is given in Table S7.3.

The refinement was carried out using only the data obtained from the long camera distance. A refinement using both the long and short data sets was attempted, but the resulting structure was much less satisfactory, with parameters refining to unfeasible values and a much poorer R factor obtained. A probable cause of this is the increased temperature used for the collection data at the short camera distance, which could have led to partial decomposition of the sample during the GED experiment. After heating the sample for data collection using the short camera distance, it was observed that a black mirror had formed on the walls of the sample tube.

Table 7.2. Refined and calculated parameters for the two-conformer SARACEN refinement of $Fe_3(CO)_{12}$. All distances are in pm and angles in degrees.[a]

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	Restraint
p_1	rFe–Fe average	262.2	257.1(13)	
p_2	rFe–Fe difference	13.8	14.2(5)	13.8(5)
p_3	rFe-C average	188.7	186.9(4)	
p_4	rFe-C difference	18.6	19.6(11)	18.6(10)
p_5	rO=C average	115.3	115.5(4)	

Parameter		B3PW91[b] (r_e)	SARACEN (r_{h1})	Restraint
p_6	rO=C difference	2.3	2.4(6)	2.4(5)
p_7	∠C-Fe(1)-Fe average	94.2	92.7(10)	
p_8	∠C-Fe(1)-Fe difference	14.2	12.4(9)	14.2(10)
p_9	∠C-Fe(2)-Fe(1) planar	175.1	171.3(9)	175.1(10)
p_{10}	∠C-Fe(2)-Fe(1)	89.6	90.4(9)	
p_{11}	∠C-Fe(2)-Fe(3)	125.5	125.6(12)	
p_{12}	∠O=C-Fe	176.5	175.4(11)	176.5(10)
p_{13}	bridge angle	66.0	57.5(21)	66.0(30)
p_{14}	∠O(bridge)-C…Fe(1)	138.5	140.9(20)	138.5(20)
Dependent parameters				
rFe(1)-Fe(2/3)		269.1	264.2	
rFe(2)-Fe(3)		255.3	250.0	
rFe-C average (terminal)		179.4	177.1	
rFe-C (bridging)		198.0	196.7	
rO=C average (terminal)		114.1	114.3	
rO=C (bridging)		116.5	116.7	

[a] For the Fe atoms, a LANL2DZ pseudopotential was used. The 6-311+G* basis set was used on all other atoms.

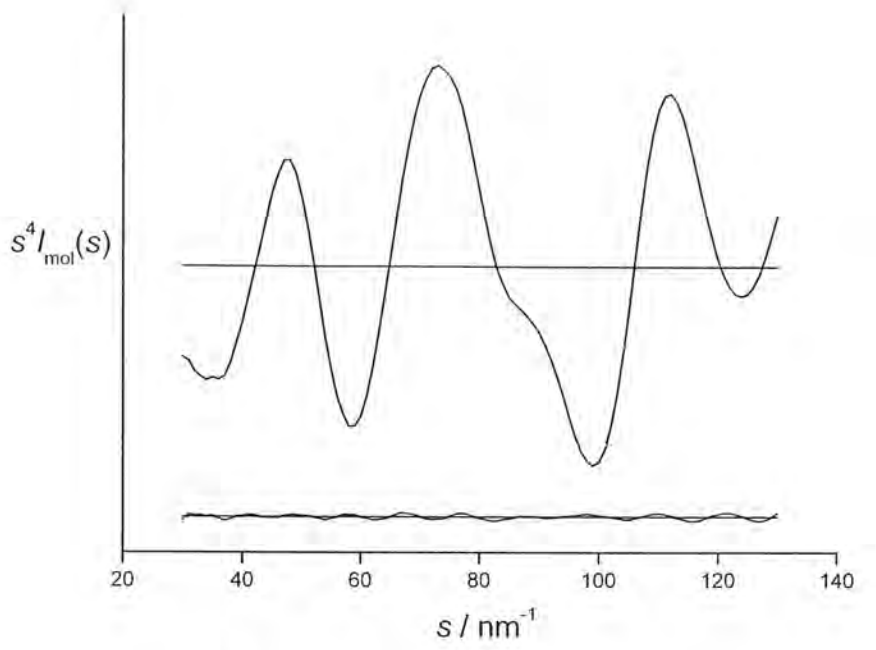


Figure 7.2. Experimental and weighted difference (experimental – theoretical) molecular scattering intensity curve for $Fe_3(CO)_{12}$.

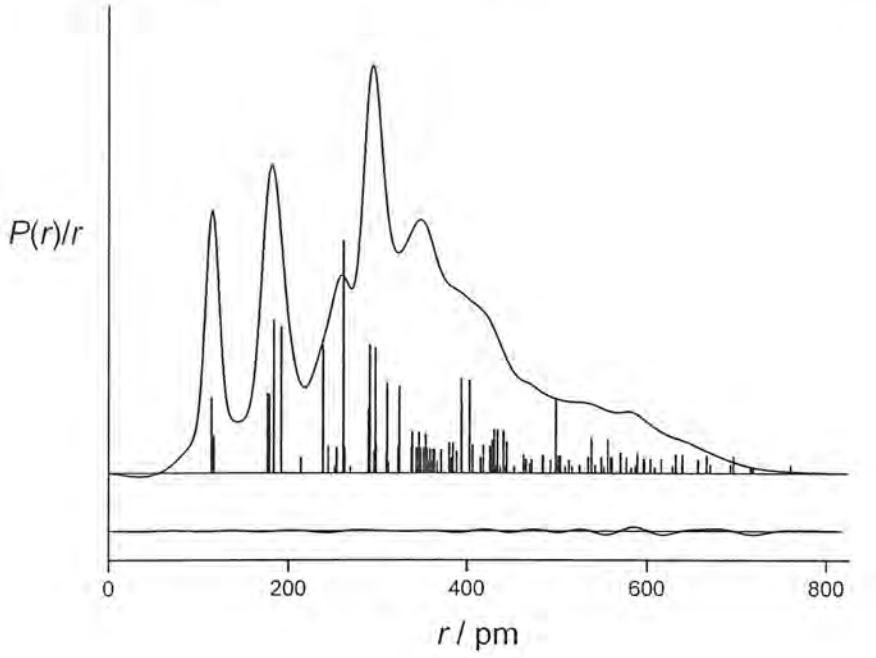


Figure 7.3. Experimental and difference (experimental – theoretical) radial distribution curves, $P(r)/r$, from the SARACEN refinement of $Fe_3(CO)_{12}$. Before Fourier inversion the data were multiplied by $s \cdot \exp(-0.00002s^2)/(Z_O - f_O)(Z_{Fe} - f_{Fe})$.

7.4 Discussion

The molecular structure of $Fe_3(CO)_{12}$ in the gas phase has been determined by GED. The refinement was carried out using the SARACEN method. It was initially intended to use the DYNAMITE and SEMTEX methods to conduct further study of the structure. However, it was discovered that the molecular mechanics package was unable to handle atoms with coordination numbers as high as those exhibited by the iron atoms in $Fe_3(CO)_{12}$, and considerable further work will be required to enable these methods to handle such molecules.

The refined gas-phase structure of $Fe_3(CO)_{12}$ shows several interesting differences when compared with the results of the theoretical calculations. The refined average distances found using SARACEN are within a few e.s.d.s of those from the B3PW91/6-311+G* calculated structure, and are generally found to be shorter for those distances involving an iron atom. For example, the average Fe-C distance is given as 186.9(4) pm using SARACEN, compared with 188.7 pm from B3PW91. This corresponds to an Fe-C distance for the bridging C atoms of 196.7 pm, The average Fe-Fe distance, which makes by far the largest contribution to the overall scattering intensity, is given as 257.1(13) pm by SARACEN, and 262.0 pm from B3PW91. The average O-C distance is in good agreement, at 115.5(4) pm from SARACEN and 115.2 pm from B3PW91. One possible cause of the comparatively large differences between theory and experiment for the distances involving an iron atom may be the use of the LANL2DZ pseudopotential to model the iron atoms in the B3PW91 calculation. However, previous theoretical work on the crystal structure of this molecule[3] has also found DFT to calculate distances involving an iron atom several picometres longer than were found experimentally.

The bridging Fe-Fe distance in the gas phase structure of $Fe_3(CO)_{12}$ is found to be considerably shorter than those in the crystal structures of both this molecule and the related triply-bridged dimer $Fe_2(CO)_9$.[10] These have been reported previously as 255.8 pm and 252.3 pm respectively, a difference of 3.5 pm. Previous mass spectrometry work on $Fe_2(CO)_9$[11] reported a bridging Fe-Fe distance of 246.0 pm, and so the value of 250.0 pm determined here for $Fe_3(CO)_{12}$ appears to follow the same trend as for the crystal structures.

In the case of the average Fe-Fe-C angle to the terminal ligands on the non-bridging Fe atom, a good agreement is seen between theory and experiment – 94.2° from B3PW91, compared with 92.7(10)° from the refinement. It was necessary to lightly restrain the difference parameter (p_8) in order for it to refine satisfactorily – this then refined to 12.4(8)°, approximately 2° lower than found computationally. The angle ∠C-Fe(2)-Fe(1) planar (p_9) is found to be less linear in the refinement than calculated – 171.3(9)°, compared with 175.1° from B3PW91. The other C-Fe-Fe angles both refined to values close to those obtained from the B3PW91 calculation: ∠C-Fe(2)-Fe(1) refined to 90.4(9) ° compared with a calculated value of 89.6°, while ∠C-Fe(2)-Fe(3) refined to 125.6(12)° compared with a calculated value of 125.5°. The parameter ∠O=C-Fe had to be restrained in order to refine satisfactorily – this then took a value of 175.4(11)°, similar to the value of 176.5° from B3PW91.

The "bridge angle" (p_{13}) describing the position of the bridging ligands relative to the plane described by the three Fe atoms had to be restrained lightly in order to refine to a sensible value, and took a value of 57.5(21)° in comparison to the value of 66.0° from the B3PW91 calculation. The parameter ∠O(bridge)-C...Fe(1) also had a loose restraint of 2.0° placed on it to ensure a satisfactory refinement. This changed from 138.5° to 140.9(20)°, suggesting the bridging O atoms remain close to planar with the Fe-Fe-C triangle.

The final goodness-of-fit parameter achieved was R_G = 0.081, showing that the refined structure of Fe_3CO_{12} matches the experimental data quite well. This is a satisfactory result considering the multiple different C=O and Fe-C distances present in the structure which are difficult to resolve accurately without the use of the SEMTEX method.

7.5 Conclusion

The molecular structure of $Fe_3(CO)_{12}$ in the gas phase has been successfully determined by GED. In keeping with theoretical predictions, a single C_{2v} – symmetric conformer is observed to give fully acceptable fit to the experimental data, with two bridging and ten terminal CO ligands. The SARACEN refinement was completed with a satisfactory goodness-of-fit to the experimental data (R_G = 0.808).

Some differences were observed between the final refined structure and that calculated at the B3PW91/6-311+G* level of theory. The most notable of these were in the angles describing the positions of the axial and equatorial C=O ligands relative to the most distant Fe atom, with the ligands in each case looking to move away from the non-bonded Fe atoms upon refinement. Overall, the result of the refinement agrees with and reinforces the theoretical picture of $Fe_3(CO)_{12}$ in the gas phase as a single doubly-bridged conformer of $C2_v$ symmetry. For future consideration, the adaptation of the SEMTEX method to deal with this and similar molecules could be expected to give better resolution of the different C-M and C=O distances and hence a better overall fit to the experimental data.

7. 6. References

1. S. Enthaler, B. Hagemann, G. Erre, K. Junge and M. Beller, *Chem. Asian J.*, 2006, **1**, 598.
2. C. Ovalles, E. Filgueras, A. Morales, C. Scott, F. Gonzalez-Giminez and B.P Embaid, *Fuel*, 2003, **82**, 887.
3. A. Cotton and J. M. Troup, *J. Am. Chem. Soc.*, 1973, **96:13**, 4155.
4. E. Hunstock, C. Mealli, M. J. Calhorda and J. Reinhold, *Inorg. Chem.*, 1999, **38**, 5053.
5. H. Chevreau, C. Martinsky, A. Sevin, C. Minot and B. Silvi, *New J. Chem.*, 2003, **27**, 1049.
6. E. Hunstock, M.J. Calhorda, P. Hirva and T. Pakkanen, *Organometallics*, 2000, **19**, 4624.
7. P. J. Hay and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 270.
8. W. R. Wadt and P. J. Hay, *J. Chem. Phys.* 1985, **82**, 284.
9. P. J. Hay, and W. R. Wadt, *J. Chem. Phys.* 1985, **82**, 299.
10. F. A. Cotton and J. M. Troup, *J. Chem. Soc., Dalton Trans.*, 1974, 800.
11. B. F. G. Johnston, J. Lewis, I. G. Williams and J. M. Wilson, *J. Chem. Soc. A*, 1967, **2**, 341-4.

8. General conclusions and future work

8.1. General conclusions

A new method of determining complex molecular structures in the gas phase has been developed. The SEMTEX (Structure Enhancement Methodology for Theory and EXperiment) method allows inclusion of data from high-level computational calculations dynamically in the refinement process, building on the previous DYNAMITE method.

The new method has been shown to be functioning correctly using the refinements of $OPBu^t_3$ and $HNPBu^t_3$ as test cases, for ease of comparison with previous SARACEN and DYNAMITE studies. As predicted, little difference in *R* factor was observed, with the peripheral atoms now adopting a set of individual parameters which mirror the results of the MP2 calculation, scaled back to the refining average parameters from GED as required.

SEMTEX (and hence DYNAMITE) have been further expanded to deal with peripheral atoms in a ring structure, using $C_6F_{11}CF_3$ as a simple test case where little difference in refinement quality would be expected between the methods. Furthermore, this example allowed the effect of the new method to be studied on refinement quality where larger peripheral atoms with greater scattering contributions are present in the molecule. The results were very satisfactory – the SEMTEX method fitted the data marginally better the SARACEN method, while the DYNAMITE method was much less successful in refining the structure.

The method has subsequently been applied to a range of other large and asymmetric molecules. The structure of $W(NBu^t)_2(NHBu^t)_2$ provided a test case for the method where more than one conformer was present, and again the SARACEN and SEMTEX refinements matched well as predicted. The studies of the $C(SiMe_2X)_4$ (X = Cl, F, Br) family of compounds provided an extremely challenging case, with multiple conformers of differing symmetry predicted and two different types of peripheral atom for the method to be applied to. Here, as hoped, the SEMTEX method showed significantly better fitting to the experimental data than SARACEN. Finally, the structure of $Fe_3(CO)_{12}$ has been determined using SARACEN, and represents a case where further development is required in order to be able to successfully apply the SEMTEX method.

8.2. Future work

The final structural refinement reported in this work, $Fe_3(CO)_{12}$, has been done using only the SARACEN method. It would be of considerable interest to see how successful SEMTEX is in handling such structures as this, where the method could be applied to either the O atoms only or (especially in more complex cases) to both the ligand atoms. To this end, of particular future interest would be the application of the SEMTEX method to the related family of $M_3(CO)_{12}$ molecules (M = Fe, Ru, Os), and also to the mixed-metal clusters $M_2M'(CO)_{12}$ (M, M' = Fe, Ru, Os), where electron diffraction data can be obtained.

Further work is required in order to make the SEMTEX (and DYNAMITE) methods applicable to structures containing atoms with such high co-ordination numbers as the M atoms in these molecules. One major difficulty is that the TINKER molecular mechanics package only handles molecules with co-ordination numbers of up to and including four. Time constraints prevented modification of this code or the adapting of DYNAMITE to use a different MM package, and this is one direction in which future work could be carried out in order to broaden further the range of molecules which can be studied using SEMTEX.

With the success of SEMTEX in the study of $C_6F_{11}CF_3$ in mind, it would be of considerable interest to see how the method performs in determining the structures of other fluorinated molecules. Other highly-fluorinated alkanes would be interesting candidates for future study, and we would again expect to see SEMTEX perform much better than DYNAMITE for these. Another possible set of candidates for future study are the family of derivatives of triflouroacetate, CF3COXR (X = O, S), several of which have recently been investigated at the University of Edinburgh.[1-5] For example, trimethylsilyl trifluoroacetate, $CF_3C(O)OSi(CH_3)_3$, contains peripheral fluorine and hydrogen atoms, each of which could affect the core $CCOSiC_3$ bond lengths and angles.

Another interesting molecule to study using the new method would be 1-trimethylsilyl-1,2,3-benzotriazole, where previous study[6] shows that the $SiMe_3$ group is co-ordinated asymmetrically to the ring system. The H-atom positions on

both the $SiMe_3$ group and the ring system could be determined with the SEMTEX method.

The family of substituted 1,3-bisketenes[7] $Me_2X[C(YMe_3)=C=O]_2$ (X,Y = Si, Ge, Sn) are known to present challenges for both theoretical and experimental structure determination, existing as multiple conformers with varying symmetries. The SEMTEX method could be used in this case to determine the positions of the peripheral hydrogen atoms and the two oxygen atoms, aiding in the complete structure determination of these molecules.

An important practical future step is to improve the automation of the SEMTEX method within the ed@ed refinement program. At present, the MP2 calculations are started manually, by taking an output co-ordinate file from the refinement when the high-level method is to be called and generating a Gaussian input file using these co-ordinates. The results of the high-level calculation must then be extracted from the output file, converted into a co-ordinate file and read back into the refinement program. Routines have been developed to convert the co-ordinate files into Gaussian input and *vice versa*, but further programming is required in order to allow ed@ed to automatically submit the high-level calculation job to a remote computer.

Once this has been realised, it would be helpful to include SEMTEX as an additional option within the existing ed@ed refinement program. Doing so would allow the structures of a broad range of large and asymmetric molecules to be confidently refined from GED data, without the need for symmetry constraints or other artificial model simplifications.

8.3. References

1. M.E. Defonsi Lestard, M.E. Tuttolomondo, E.L. Varetti, D.A. Wann, H.E. Robertson, D.W.H. Rankin and A. Ben Altabef, *J. Mol. Struct.* **917**, 2009, 183.
2. M.E. Defonsi Lestard, M.E. Tuttolomondo, E.L. Varetti, D.A. Wann, H.E. Robertson, D.W.H. Rankin and A. Ben Altabef, *J. Raman Spectrosc.*, 2009, doi:10.1002/jrs.2371.
3. M.E. Defonsi Lestard, M.E. Tuttolomondo, D.A. Wann, H.E. Robertson, D.W.H. Rankin and A. Ben Altabef, *J. Raman Spectrosc.* 2010, in press.
4. M.E. Defonsi Lestard, M.E. Tuttolomondo, D.A. Wann, H.E. Robertson, D.W.H. Rankin and A. Ben Altabef, *J. Chem. Phys.* **131**, 2009, 214303-1.
5. M. E. Defonsi Lestard, M. E. Tuttolomondo, E. L. Varetti, D. A. Wann, H. E. Robertson, D. W. H. Rankin and A. Ben Altabef, *J. Mol. Struct.*, 2010, in press.
6. T. Foerster, D. A. Wann, H. E. Robertson and D. W. H. Rankin, *Dalton Trans.*, 2009, 3026-3033
7. K. B. Borisenko, R. N. Yezhov, S. V. Gruener, H. E. Robertson and D. W. H. Rankin, *Inorg. Chim. Acta*, 2008, 467-472.

Appendix A

Conferences and courses attended

Conferences

21st Austin Symposium on Molecular Structure
Austin, Texas, USA March 5-8 2006
Poster presentation: *Combining theory and experiment: enhanced methods for the dynamic interaction of theory and experiment*

USIC (Universities of Scotland Inorganic Conference)
St Andrews, September 2006
Poster presentation: *Combining theory and experiment: enhanced methods for the dynamic interaction of theory and experiment*

ScotCHEM Computational Chemistry Symposium
St Andrews, 4th April 2007
Poster presentation: *Combining theory and experiment: enhanced methods for the dynamic interaction of theory and experiment*

12th European Symposium on Structural Chemistry
Bleubeuren, Germany June 25-28 2007
Poster presentation: *Dynamic combination of ab initio and gas electron diffraction data – the SEMTEX method*

22nd Austin Symposium on Molecular Structure
Austin, Texas, USA March 4-7 2008
Poster presentation: *The SEMTEX method of total gas-phase structure determination*

Courses

Introduction to UNIX, 2005

Introduction to FORTRAN, 2005

UNIX 2: Enhancing your UNIX skills, 2006

Molecular Modelling 4 Chemists, Cardiff, 2006

Departmental colloquia, 2005-2008

Inorganic section meetings, 2007-2008

Materials, Structure and Chemical Physics section talks, 2005-2008

Appendix B

Publications

The following pages are print-outs of published articles arising in whole or in significant part from this work. The results obtained by myself and relating to this thesis are reported in Chapters 2 and 4.

1. **Structure enhancement methodology using theory and experiment: gas-phase molecular structures using a dynamic interaction between electron diffraction, molecular mechanics, and ab initio data.** G. R. Kafka, S. L. Masters and D. W. H. Rankin., *J. Phys. Chem. A*, 2007, **111**, 5913.
2. **Structural tungsten-imido chemistry: the gas-phase structure of $W(NBu^t)_2(NHBu^t)_2$ and the solid-state structures of novel heterobimetallic W/N/M (M = Rh, Pd, Zn) species.** H. Choujaa, S. D. Cosham, A. L. Johnson, G. R. Kafka, M. F. Mahon, S. L. Masters, K. C. Molloy, D. W. H. Rankin, H. E. Robertson and D. A. Wann., *Inorg. Chem.*, 2009, **48**, 2289.